Canada College Learning Center

Math Jam
Statistics

Student Handbook

Written by Kazumi Tsuchiyose
Professor of Mathematics
College of San Mateo
1700 Hillsdale Boulevard
San Mateo, CA 94402

Contents

Math Jam Day 1

Agenda for day 1

Introductions

Introductory meeting at 9:30

Student agreements

Do you have a My Stat Lab account? Can you log in?

Do you have a @my.smccd.edu Canada College student email account?

Do you have a Canvas account and can you log in?

Please take the pre-test. Note your score and identified areas of weakness

Plans for the week

Note to Dave - write a page on formulas and definitions

Terms, definitions, formulas, and concepts

- mean
- mode
- median
- midrange
- standard deviation

- variance
- normal curve
- even distribution
- right skew
- left skew
- weighted average
- z-score
- qualitative and quantitative data
- nominal and ordinal data

Math Review

1. On a number line please graph:

a) $X > 1$

b) $X < 2$

c) $-1 < X < 2$

d) $X < -1$ or $X > 1$

e) $-2 < X < 2$

f) $X > -1$

2. X = 4,3,6,5,2

a) Make a dot plot
b) Compute ∑X
c) Compute ∑(X-2)
d) Compute ∑X^2
e) Compute ∑(X-2)^2
f) Compute ∑X/n

3. Name examples of an experiment with exactly two observable outcomes. A binomial experiment yields a binomial distribution.

4. How many different ways (boy or girl) are there to have one child? Two children? Three children?

5. Graph y = 2x + 1. Give some examples of two variables that have a linear relationship.

6. Give any value of x between 0 and 1 (inclusive) and give an example of events with a probability between 0 and 1 (inclusive). 0 means no probability and 1 means the event always happens.

7. There are ___ male students in this class of ___ students. What is n? What is x?. What is the portion (%p) of male students?

Vocabulary and terminology

8. Definitions

- Population
- Variable
- Data
- Parameter
- Sample
- Statistic
- Experiment
- Observational study
- Statistic
- Voluntary sampling
- Convenience sampling
- Categorical data
- Quantitative data
- Numerical data
- Qualitative data

- Continuous data
- Discrete data

9. A drug manufacturer is interested in the proportion of persons who have hypertension whose condition can be controlled by a new drug. A study of 5,000 people with hypertension is conducted and it is found that 80% of the individuals are able to control their hypertension with the drug. Assuming 5,000 individuals are representative of the group who has hypertension answer the following:

- What is the population?
- What is the sample?
- Identify the parameter of interest
- Identify the statistic and give its value?
- Do you know the value of the parameter?

10. Consider the following sample: 7,6,10,7,5,9,3,7,5,13

Make a dot plot. What is the:

- mean
- mode
- median
- midrange

Using the TI-83 and TI-84 Calculator

11. Consider the following sets of exam scores:

a) 47,79,79,82,83,85,86,87,89,90,91,96

b) 31,35,39,55,56,63,63,64,67,68,78,83

Generate the five number summary on a TI-83. What is the:

- n
- min
- Q1
- Q2
- Median
- Q3
- max

Generate a box plot on the TI-83

12. Interpreting histograms. Describe overall shape of:

- Bell
- Uniform
- Skewed right
- Skewed left

13. Examples of statistics

- Jenny Craig weight loss statistics
- What is that fish on your plate?
- United Airlines overbooking

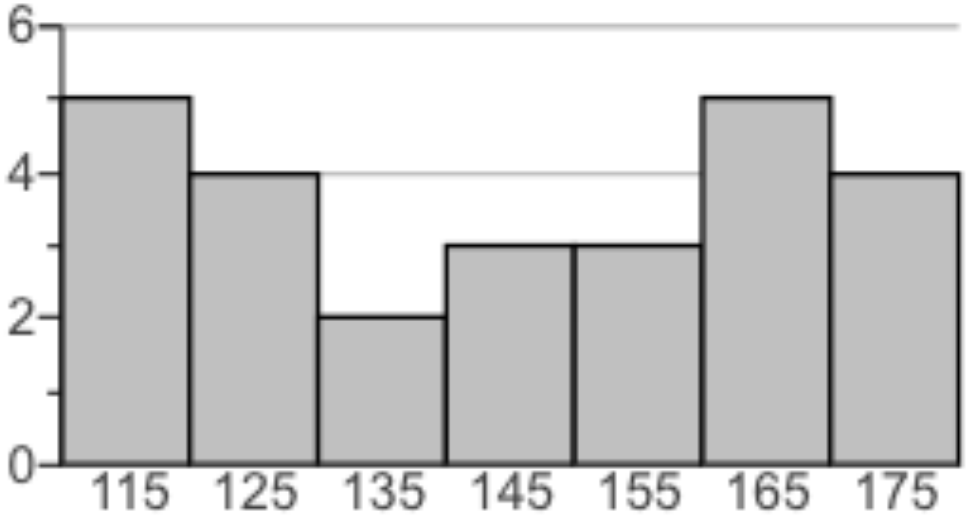
Example of conditional probability

- Sixty eight percent of Whole Foods shoppers are Amazon Prime members.
- Incomes by college degree

Instructor notes

- Histogram
- Many normal distributions
- Graph
- X,Y coordinate graph
- Chapters 1 and 2 pre-test
- Shade middle 50 percent, top two percent, bottom 90 percent, probability greater than .90
- Some of the inequality in a Normal Distribution graph

14. This histogram shows distribution of exam scores for 26 students in a biology class.



- How would you describe the shape of this distribution of exam scores? Use course vocabulary.
- Give an interval that describes typical grades on this exam.
- Estimate the overall range of grades on this exam. Range = Max - Min.
- What percentage of the students made a D on the exam? This is a grade of 60 - 69%
- What percentage of the students passed the exam with a 70 or better?
- What percentage of the students made an A or B?
- What percentage of the students who passed the exam made an A or B?
- What percentage of students who failed the exam with a grade lower than 70 made a D grade of 60 to 69 percent.

Chapter 1 Pre-Test

1. An author wrote a book about bilingualism in America. Her conclusions were based on 5,212 replies received after mailing 123,628 questionnaires. Are the conclusions likely to be valid in the sense that they can be applied to the general population? Why or why not?

____ Yes, enough responses were received to make valid conclusions

____ Yes, the numbers were precise

____ No, not enough responses were received to make valid conclusions

____ No, it was a voluntary response sample

2. In the study of a weight loss program, 30 subjects lost a mean of 3.4 lbs after 12 months. Methods of statistics can be used to show that if this diet had no effect the likelihood of getting these results is three chances in 1,000.

Does the weight loss program have statistical significance?

____ Yes because the results are unlikely to occur by chance

____ No because the results are likely to occur by chance

____ No because the results are unlikely to occur by chance

____ Yes because the results are likely to occur by chance

Does the weight loss program have practical significance?

____ No the results have a chance to occur even if the weight loss program had no effect

____ Yes the results are too unlikely to occur by chance

____ No someone starting a weight loss program would likely want to lose considerably more than 3.4 lbs

____ Yes the amount of lost weight is large enough to be considered practically significant

3. Determine whether the given value is a statistic or a parameter. A sample of students is selected and is found that 45% own a computer. Choose the correct statement below.

____ Parameter because the value is a numerical measurement describing a characteristic of a population

____ Statistic because the value is a numerical measurement describing a characteristic of a sample

____ Statistic because the value is a numerical measurement describing a characteristic of a population

____ Parameter because the value is a numerical measurement describing a characteristic of a sample

4a. Convert the fraction 7/10 to an equivalent percentage

4b. Convert 53.3% to an equivalent decimal

4c. What is 96% of 200?

4d. Convert 0.144 to an equivalent percentage

5. Use common sense to determine whether the given event is either impossible, possible but very unlikely, or possible and likely. While driving through a city you arrive at three consecutive traffic lights and they are all red. Choose the correct answer below:

____ Possible and likely

____ Possible but very unlikely

____ Impossible

6. Based on a study of heights of men and women who play basketball a researcher concludes that the exercise from playing basketball causes people to grow taller. Do you agree with this conclusion? Choose which conclusion you agree with:

____ No there is no relationship between height and playing basketball

____ Yes people who play basketball need to be tall so they have evolved with a height advantage

____ No there may be a relationship between height and playing basketball but that does not mean that one causes the other

____ Yes people who exercise grow stronger bones and are more likely to grow taller

7. Determine whether the following value is from a discrete or continuous data set: Amount of sugar in a bowl is 0.75 cups. Is this value from a discrete or continuous data set?

____ Discrete

____ Continuous

8. Determine whether the description corresponds to an observational study or an experiment. Research is conducted to determine if there is a relationship between heart arrhythmias and caffeine consumption. Does this description correspond to an observational study or an experiment?

____ Experiment

____ Observational study

Math Jam Day 2

Math Review

1. Given X = 2,9,3,6,5 compute the following

a) ∑X

b) ∑X/n

c) ∑X^2

d) ∑(X-5)

e) ∑(X-5)^2

f) ∑(X-5)^2/(n-1)

g) Xmax – Xmin

Combinations

2. What are the possible gender combinations in a family, if there are:

- Two children
- Three children
- Four children
- Five children
- Six children

3. What is an experiment with two outcomes? What are examples of activities with an observable outcome?

4. The mean checkout time for all customers at Costco is to be estimated by using the mean checkout time obtained by 75 randomly selected customers. Match the items in column II with the statistical term in column I.

I	II
Data(one)	a) Process used to select the 75 customers and measure their check out time
Data (set)	b) The 75 customers
Parameter	c) The mean time for all customers
Statistic	d) 2 mins, one customers checkout time
Population	e) All customers at Costco
Sample	f) Checkout time or one customer
Variable	g) 75 times
experiment	h) Mean time for 75 customers

Descriptive Statistics Terminology

5. Distribution terminology

- Center, cluster, mode
- Spread, outlier, range, gaps
- Shape, overall pattern
- Bell shape
- Left or right skew
- Uniform
- Symmetric

6. Weighted mean problem

GPA : A (5 units), B (4 units), C (6 units), A (3 units), B (3 units)

Compute grand mean from 2 classes:

- Class A has 30 students and the mean test score is 70 pts.
- Class B has 40 students and the mean test score is 80 pts.

7. Complement is from 100% total. What is the complement of:

- 95%
- 68 %
- 99,7%
- 50%

8. What are these symbols? Write the definition of each symbol with math notation:

- S
- S^2

- σ
- σ^2
- μ
- $\bar{x}$
- $\hat{p}$

9. Five number summary definitions:

- Min
- Q1
- Q2 (aka Median)
- Q3
- Max
- IQR = Q3 - Q1
- Outliers

10. A study of small business failures looked at 148 food and drink businesses in central Indiana. 106 were headed by men and 15 failed within 3 years. 42 were headed by women and 7 failed within 3 years.

- What portion of the business failed given they are headed by women?
- Make a two way table.

12. Example of statistics.
    - and and or conditional probability
    - two way table

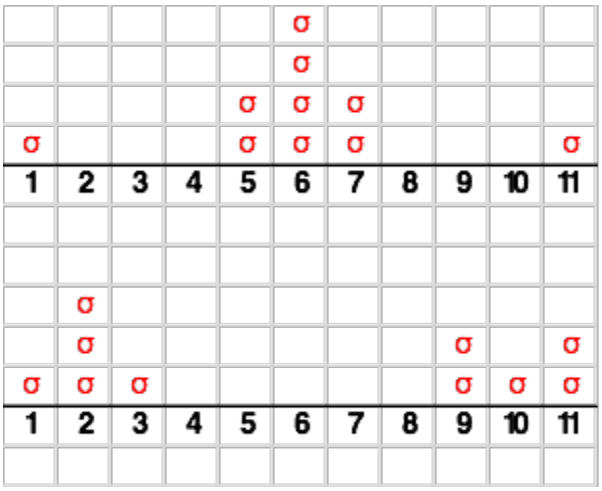
13. Frequency Distribution
    - Range
    - number of classes
    - class width
    - starting point
    - add class width to list
    - lower class limit
    - histogram. Please note that a histogram has no gaps

14. Given the following test data:

X(age)	Frequency	Relative Frequency
16	3	
17	2	
18	10	
19	9	
20	5	
21	1	

- Fill in relative frequency using the fraction.
- What is n?
- What is $\sum(f(i)/n)$?
- Compute sum of relative frequency
- What is the mean age from this group?
- Population mean μ = ____

15. Look at these two dot plots



Which data set has a standard deviation of 2.45, and which data set has a standard deviation of 4.29

16. More definitions:

- Empirical rule
- Relative standing
- Z-score
- Usual value.
- Estimate of s by range /4

17. What is the coefficient of variation?

- For a sample: $CV = s/\bar{x}$
- For a population: $CV = \sigma/\mu$

18. Which of the following is more extreme?

- Height of men $x = 76.2$in ; $\mu = 68.3$ in, $\sigma = 3.02$ in,
- Weight of men $x = 237.1$ Lb; ; $\mu = 172.6$ Lb, $\sigma = 26.3$ Lb

Use the Z-score to answer this question.

19. If a student football player has a z-score of 4 from his weight, what is his weight?

20. IQ of Einstein $x = 150$; $\mu = 100$, $\sigma = 15$. Compute the Z-score

Chapter 2 Pre-Test

The frequency distribution below represents frequencies of actual low temperatures recorded during the course of a 31-day month. Use the frequency distribution to construct a histogram.

Does the data appear to have a distribution that is approximately normal?

____ Yes it is approximately normal

____ No it is approximately uniform

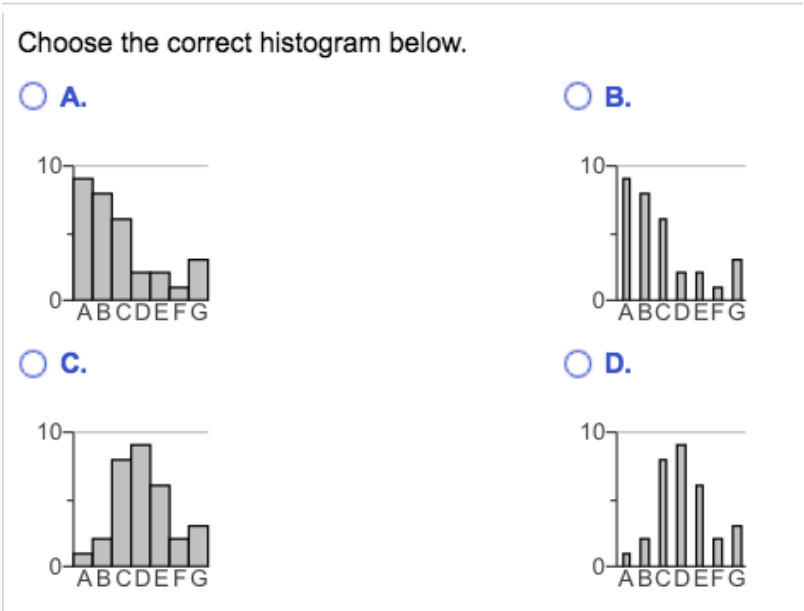
____ No it is not at all symmetric

2. From the table below identify the class width, class midpoints, and class boundaries.

- What is the class width?
- In ascending order, what are the
  - class midpoints?
  - class boundaries?

Class	Frequency
A 39 - 44	1
B 45 - 50	2
C 51 - 56	5
D 57 - 62	9
E 63 - 68	8
F 69 - 74	3
G 75 - 80	2

Choose the correct histogram



3. According to a recent study women have an average (mean) height of 57 inches or 145 centimeters, and men have an average height of 66 inches, or 168 centimeters. These averages are shown in the bar graph below. Do these graphs depict the data correctly?

57 inches, or 145 centimeters

66 Inches, or 168 centimeters

____ No because the heights are one dimensional measurements but the graph uses objects of area

____ Yes because the data are represented in both standard and metric units

____ No because the graph contains errors

____ Yes because the heights are two dimensional measurements, it is appropriate to depict the data with objects of area

4. Numbers in the table below represent average daily rainfall for one month.

0.41	0	0	0.28	0	0.42
0	0.22	0	0	1.33	0
0.22	0	0.02	0	0.15	0
0.07	0.44	0	0.02	0	0.22
0	0.08	0	0	0.06	0

•Construct a frequency distribution beginning with a lower class limit of 0.00 and use a class width of 0.20.

Daily Rainfall in Inches	Frequency	Daily Rainfall In Inches	Frequency
0.00 - 0.19		0.80 - 0.99	
0.20 - 0.39		1.00 - 1.19	
0.40 - 0.59		1.20 - 1.39	
0.60 - 0.79			

- Does the frequency distribution appear to be roughly a normal distribution?

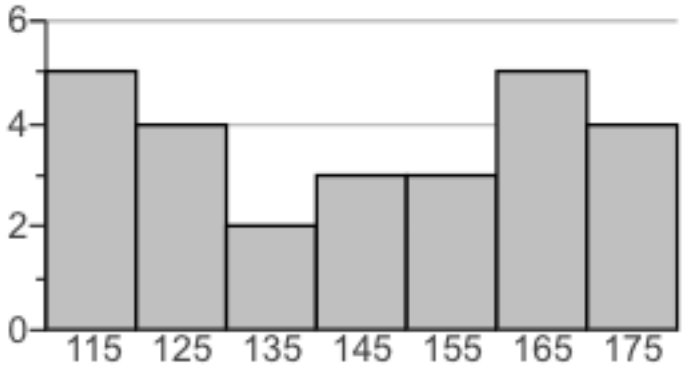
  ____ Yes all of the requirements are met

  ____ No although the frequencies start low, increase to some maximum, then decrease, the distribution is not symmetric

  ____ No the distribution is not symmetric and the frequencies do not start off low

  ____ No although the distribution is approximately symmetric, the frequencies do not start low, increase to some maximum frequency

5. The histogram to the right represents weight in pounds of a certain high-school math team.



How many team members are represented in this histogram?

Math Jam Day 3

Interpreting Two-Way Tables

1. How many results in the following table are false positives? False negatives? Do you think the test is accurate?

	No Cancer	Yes Cancer	Total
Blood test (+)	16	45	
Blood Test (-)	32	13	
total			

2. Rangers at Yosemite took a sample of 6 rats to see if they carry a hantavirus. How many different results are possible? How likely do they see three rats out of six have this virus? Can you do it without listing them all? Can you use the notation of C for each case?

3. Do you think venture capital companies have a disproportionately smaller number female CEOs than other companies?

	Male	Female	Total
Venture Capital Company		3	100
Not a Venture Capital Company		23	500
total			

4. What is the range of the "usual" height for men if $\mu = 68$ inches and $\sigma = 3$?

5. What is the area of the rectangle with:

- L = 4 in, W = 2 in
- L = 5 in W = 2 in
- L = 5 in, W = .2 in
- L = 6 in , W = 1/6 in

Definitions

6. Define distribution

7. What are these symbols and terms? Write the definition of each. Use a formula and math notation when appropriate

- S
- S^2
- σ
- σ^2

- μ
- $\bar{x}$
- z
- CV
- M
- Q1
- Q2
- Q3
- IQR
- Range
- Mid range
- n
- x
- p

7. What is the difference between these two representations of the coefficient of variation:

- $CV = s/\bar{x}$
- $CV = \sigma/\mu$

Understanding Two Way Tables

8. The Hawk-eye instant replay system for tennis was introduced in the US open.

- Men challenged 489 referee calls and 201 of them were successfully upheld by the Hawk-eye system.
- Women challenged 350, and 126 of them were successful

Fill in the table. For each gender, what is n, x, and p.

	Male	Female	Total
Successful challenge			
Unsuccessful challenge			
Totals			

9. What is the complement event of at least one student will pass this class.

10. Given the probability of having a girl is .5, compute the probabilities of 0, 1, 2, 3, and 4 girls in a family with four children.

11. Using “and” and “or” logic with a two-way table

	Male	Female	Total
Yes I have a tatoo	18	13	
No I do not have a tatoo	119	55	
Totals			

From the above table compute the following probabilities:

$\hat{p}(N)$ $\hat{p}(M) =$

$\hat{p}(Y)$ $\hat{p}(F) =$

$\hat{p}(F \text{ or } Y)$ $\hat{p}(F \text{ or } N) =$

$\hat{p}(F \text{ and } Y)$ $\hat{p}(F \text{ and } N) =$

12. TI-83/84 calculator exercise.

- Enter the “x” values as List L1 in the TI-83/84 calculator.
- Enter the “y” values as List L2.
- Generate scatter plot from the data using the calculator

x	10	20	30	40	50
y	11	19	28	43	50

- Is there any correlation between x and y?
- Is it a linear correlation?

Probability

13. Disjoint events are mutually exclusive events. These are events that can not happen at the same time. What are examples of two events that cannot happen at the same time?

14. What is the probability of getting “1” and “1” on two dice when two dice are rolled? Is it unusual to get two 1s? Why or why not?

15. If the probability of a student who took Math Jam passing a test is 0.9, then what is the probability the student will not pass the test?

16. The probability of the alarm clock working is 0.95. If two identical alarm clocks are set to make sure a student will wake up in the morning for a final exam at 8:00 am, what is the probability that both alarms will fail?

17. Examples of independent events

- Pick a card from a deck. Pick the second card *with* replacement
- Pick a card from a deck. Pick the second card *without* replacement

18. What is the quantity $x - \bar{x}$ called?

19. Probability terminology. Define the following:

- Event
- Activity
- Sample space

20. Two coins are tossed. The number of possible outcomes is ____

21. How many two digit numbers can be formed using each of the digits 3, 5, and 7 if each number is used only once ____

22. A box contains four slips of paper. Each slip has one of the letters g, a, t, or e written on it. The number of three letter outcomes is ____

23. From six students, four boys and two girls, one boy and one girl is selected. The number of possible different outcomes is ____

24. A girl has five blouses and three pairs of slacks. How many different outfits of one blouse and one pair of slacks are possible ____

25. The number of different ways five students can be arranged in a row is ____

26. Given four digits 2, 5, 6, and 8, how many different arrangements of these four numbers is possible if each digit is used only once ____

27. A school cafeteria offers six different sandwiches and four different drinks. How many combinations of one drink and one sandwich are possible ____

28. Two dice are tossed. The number of ways eight can appear as the sum is ____

29. The probability of an event happening is 3/7. What is the probability of the event not happening ____

30. If the probability of a candidate winning an election is 0.6 then what is the probability of that candidate not winning the election ____

31. Mr. Lapuz bought 12 chances on a car raffle. If 6,000 chances are sold what is the probability he will now win the car ____

32. If two coins are tossed what is the probability of one head and one tail ____

33. Two dice are tossed. The probability of obtaining a sum of 12 is ____

34. A card is drawn from a deck of 52 cards. What is the probability of drawing a heart ____

35. There are three red buttons in a bag, along with two greens and five yellows. One button is drawn. The probability that a red button is drawn is ____

36. One letter of the name Julane is selected at random. The probability that the selected letter is a vowel is ____

37. One card is drawn from a deck of 52. What is the probability that the card is a jack, queen, or king ____

38. A jar contains two red marbles, four blue, and two yellow. If a single marble is selected at random what is the probability that a red or yellow marble is chosen ____

Chapter 3 Pre-Test

1. Statistics are sometimes used to compare or identify authors of different works. Lengths of the first ten words in books by authors Terry and David are listed below. Find the mean and median for each of the two samples and then compare the two sets of results.

Terry: 3 3 5 3 3 9 1 2 2 7

David: 3 5 3 3 3 2 3 4 2 4

The mean number of letters per word in Terry's book is: ____

The mean number of letters per word in David's book is: ____

The median number of letters per word in Terry's book is: ____

The median number of letters per word in David's book is: ____

Compare the two sets of results. Does there appear to be a difference:

____ Yes. Based on the results words in Terry's book are longer than words in David's book

___ No. Based on the results, words in Terry's book are the same length as words in David's book

___ Yes. Based on the results, words in Terry's book are shorter than the words in David's book

2. A simple random sample of pages from a particular dictionary was obtained. Listed below are the numbers of words defined on those pages. Find the (a) mean, (b) median, (c) mode, (d) midrange for the given sample data. Given that this dictionary has 1,409 pages with defined words, estimate the total number of defined words in the dictionary. Is that estimate ikely to be an accurate estimate of the number of words in the English language?

48 68 34 42 33 57 77 40 54 74

The mean number of words per page is: ____

The median number of words per page is: ____

The mode is: ____

The midrange is: ____

The estimated number of words in the dictionary is: ____

Is the estimate likely to be an accurate estimate of the number of words in the English language?

____ Yes the sample mean is always an accurate estimate of the population mean regardless of sample size.

____ Yes because the mean is sensitive to extreme variations in sample values

____ No because the sample in this problem is too small to accurately represent the population

____ No it is an accurate estimate for the total number of defined words in the dictionary but it is likely that the dictionary does not contain every word in the English language

3. Below are 36 sorted ages of an acting award winner. Find P(40) using the method represented in the text book.

16 17 17 19 20 22 22 27 27 27

30 30 31 31 32 32 34 34 35 35

36 38 38 38 41 43 46 46 52 55

55 57 63 63 64 74

P(40) = ____

4. With a height of 61 inches Arthur was the shortest president of the statistics club in the past century. Club presidents of the past century have a mean height of 69.7 inches and a standard deviation of 1.2 inches.

Compute the following:

Positive difference between Arthur's height and the mean is: ____ inches.

That difference is ____ standard deviations

The z score is ____

Is Arthur's height unusual?

5. Statistics students participated in an experiment to test their ability to determine when one minute (sixty seconds) has passed. Results below are in seconds.

52 53 60 67 54

Compute the following:

Range ____

Sample variance ____

Sample standard deviation ____

Identify one reason why the standard deviation from this sample might not be a good estimate of the standard deviation for the population of adults

____ The standard deviation may not be a good estimate because htere is most ikely an outlier in the sample data

____ The standard deviation may may not be a good estimate because the data are from a random sample

____ The standard deviation may not be a good estimate because the sample is very small

____ The standard deviation may not be a good estimate because the standard deviation is too large

6. Listed below are the playing times in seconds of songs that were popular at the time of this writing.

451 239 229 248 248 297 276 226

247 213 264 235 213 255 250 259

Compute the following:

mean ____

median ____

mode ____

midrange ____

Is there one time that is very different from all the others?

____ Yes the time of 451 seconds is very different from all the others

____ No all the times are not very different from each other

____ Yes the time of 297 seconds is very different from the others

____ Yes the time of 297 seconds is very different from the others

____ Yes the time of 213 seconds is very different from the others

7. Compute the mean temperature from the data summarized in the following frequency distribution. Compare the computed mean to the actual mean of 51.8 degrees Fahrenheit

Fahrenheit Temperature	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64
Frequency	3	6	13	6	1

Mean of the frequency distribution is: ____

Which of the following best describes the relationship between the computed mean and the actual mean:

____ The computed mean is not close to the actual mean because the difference between the means is less than 5%

____ The computed mean is close to the actual mean because the difference between the means is less than 5%

____ The computed mean is not close to the actual mean because the difference between the means is more than 5%

____ The computed mean is close to the actual mean because the difference between the means is more than 5%

8. Listed in the table below are costs in dollars of round trip flights between two cities. Find the coefficient of variation for each of the two sets of numbers and compare the variation. Round all answers to one decimal place as needed. The coefficient of variation is the standard deviation divided by the mean.

30 days in advance	290	243	263	241	311	240	256
1 day in advance	455	625	551	904	625	1088	544

The coefficient of variation for prices of tickets purchased 30 days in advance is: ____

The coefficient of variation for prices of tickets purchased one day in advance is: ____

Is there a difference in variation between the two data sets?

____ The costs of tickets purchased 30 days in advance have less variation than the costs of tickets purchased one day in advance

____ There is no significant difference in the variation

____ The costs of tickets purchased one day in advance have less variation than the costs of tickets purchased 30 days in advance

Math Jam Day 4

Probability

1. The following table shows the number of people who died in a shipwreck, by the class of cabin they were booked in.

	Died (D)	Survived (S)	Total
Upper class (U)	16	126	
Middle class (M)	13	90	
Lower class (L)	107	101	
Total			

What is the meaning of the following probabilities:

- $\hat{p}(S)$
- $\hat{p}(S/U)$
- $\hat{p}(S \text{ and } U$
- Compare $\hat{p}(S)$ and $\hat{p}(S/U)$. What can you tell about surviving this shipwreck?

If you want to see if D and L are independent, what probability do you need to compare?

If the category of cabin made no difference in survival rates what would numbers in the following table look like?

	Died (D)	Survived (S)	Total
Upper class (U)			60
Middle class (M)			90
Lower class (L)			150
Total	100	200	300

2. Given the following table determine the following probabilities

	Female	Male	Total
Yes I have a tatoo	18	13	
No I do not have a tatoo	119	55	
Total			

P(____)= 68/205 P(____)=13/68 P(____)=13/205

P(____)=174/205 P(____)=119/174 P(____)=119/204

P(____)=31/205 P(____)=18/31 P(____)=18/205

3. What is the complement event of: At least nine students did not take the Ch 2 test.

4. If X is the number of girls among 4 siblings:

- $\hat{p}(X=0)=$
- $\hat{p}(X=1) =$
- $\hat{p}(X=2) =$
- $\hat{p}(X=3) =$
- $\hat{p}(X=4) =$

5. Given a quiz with two true/false questions. What is the probability of getting the first and the second question correct using pure guesses. What is the probability of getting both questions correct if you studied hard?

6. The probability of a student passing a Math 200 test is 0.6. What is the probability of failing the test? What are the odds agains passing the class? What are the odds of not passing the class?

7. The probability of an airplane engine functioning is .98. If there are two of the identical engines are on the plane, how likely do you die from crashing with engine failure?

8. What is the quantity $x - \bar{x}$ called?

9. Given a frequency table what is the sum of all relative frequencies? What is the sum of all absolute frequencies?

Using the TI-83 and TI-84 Calculator

10. Using the TI-83 or 84 calculator, enter the following numbers into L! and L2 ad then generate a scatter plot.

Budget (x)	61	88	51	36	204	96	90
Gross (y)	63	62	49	53	515	142	52

Is there any correlation between x and y? Is it a linear correlation? If so, can you model the scatter plot by a linear equation?

12. Which of the following two tests did a student do better in?

- English test sore 80, class mean is 75, and class standard deviation is 5
- Math test score is 75, the class mean is 60 and the class standard deviation is 10

My notes

- Ch 5 pre test, practice binomialpdf, cdf, at least, normal cdf
- Expected value, geogebra, usual values, margin of errors, P value= probability value

Chapter 4 Pre-Test

1. When fifteen job applicants are examined for a certain ability, at least one of them tests negative.

2. Which of the following values are valid probabilities and which are not?

- 5/3
- 3/5
- 0
- -0.45
- 1
- 1.58
- 2^(1/2)
- Square root of two
- 0.04

2. Determine the written complement of the following event.

When fifteen job applicants are examined for a certain ability, at least one of them tests negative.

Which of the following is the complement

____ No applicants test negative

____ All applicants test negative

____ No applicants test positive

____ More than one applicant tests positive

3. It is impossible to get five aces when selecting cards from a shuffled deck. Express the indicated likelihood as a probability value between 0 and 1 inclusive.

The probability of getting five aces when selecting cards from a shuffled deck is ____

4. Numbers in the following table summarize blood groups and Rh types for 100 typical people.

	Group O	Group A	Group B	Group AB
Rh+	35	30	12	3
Rh-	9	8	1	2

If one person is randomly selected what is the probability of getting someone who is in group A and Rh-?

5. Evaluate the given expression and express the result using the usual format for writing numbers.

- 2!=
- 3!=
- 10!=
- 23!=

6. For each given pair of events, classify the events as independent or dependent. If two events are technically dependent but can be treated as independent then the events are independent.

a. A person finds the microwave is not working and then finds the kitchen light is not working. Choose the correct answer below:

____ The events are dependent since the microwave not working may increase the probability of the kitchen light not working.

____ The events are independent since the kitchen light not working has no effect on the probability of the microwave not working

____ The events are dependent since the microwave not working has no effect on the probability of the kitchen light not working

____ The events are independent since the microwave not working has no effect on the probability of the kitchen light not working

b. Randomly selecting a boy at the mall, and then randomly selecting a second boy at the mall. Choose the correct answer below.

____ The two events are technically dependent but can be treated as independent because small samples are drawn from a large population

____ The events are independent since randomly selecting a boy at the mall has no effect on the probability of randomly selecting a second

____ The events are dependent since randomly selecting a boy at the mall may decrease the possibility of randomly selecting a second

c. The habit of driving recklessly and the probability of getting into a serious car accident. Choose the correct answer below:

____ The probability of getting into a serious car accident could affect the habit of driving recklessly so the events are dependent

____ The habit of driving recklessly could affect the probability of getting into a serious accident, so the events are independent

____ The habit of driving recklessly could not affect the probability of getting into a serious car accident so the events are independent

____ The habit of driving recklessly could affect the probability of getting into a serious car accident so the events are dependent

7. To the right are outcomes that are possible when a couple has three children. Refer to the list and compute the following probabilities:

a. Zero boys

b. Two boys

c. Three girls

d. One girl

e. One boy

f. Zero girls

8. If a couple plans to have four children what is the probability there will be at least one boy?

- The probability is: ____

Can the couple be confident of having at least one boy?

____ Yes because the probability is close to 1

____ No because the probability is close to 0

____ No because the probability is close to 1

____ Yes because the probability is close to 0

9. Decide whether the following two events are disjoint:

- Randomly selecting someone who owns a car, and

- Randomly selecting a married male

Are the two events disjoint?

____ No because the events cannot occur at the same time

____ Yes because the events can occur at the same time

____ Yes because the events cannot occur at the same time

____ No because the events can occur at the same time

19. Answer the following questions

a. If P(A) = 0.05 then what is the complement of P(A)

b. If a certain group of women has a 0.66% rate of red and green color blindness. If a woman is randomly selected what is the probability she does not have red and green color blindness?

Chapter 5 Pre-Test

1. Determine whether the procedure described below results in a binomial distribution. If it is not binomial identify at least one requirement that is not satisfied

Five hundred different voters in a region with two major political parties, A and B, are selected randomly from the population of 4.3 million voters. Each voter is asked if they belong to party A or B

Choose the correct answer below:

____ No there are more than two possible outcomes

____ No the probability of success is not the same in all trials

____ No the number of trials is not fixed

____ Yes the result is a binomial probability distribution

____ No the trials are not independent and the sample is more than five percent of the population

2. Based on data from a car bumper sticker study, when a car is randomly selected, the number of bumper stickers and the corresponding probabilities are as shown to the right.

a. Does this information describe a probability distribution?

b. Assuming that the numbers reflect a probability distribution, please compute:

____ Mean

____ Standard deviation

c. Use the range rule of thumb to identify the range of usual values.

____ Minimum usual value

____ Maximum usual value

d. Is it usual for a car to have more than one bumper sticker?

____ No because the probability of more than one bumper sticker is 0.123 which is greater than 0.05

____ No because the probability of having one bumber sticker is 0.077 which is greater than 0.05

____ Yes because the probabilities for random variable x from 2 to 9 are all less than 0.05

____ Not enough information is given

3. Determine whether the following probability experiment represents a binomial experiment and explain the reason for your answer.

An experimental drug is administered to 200 randomly selected individuals, with the number of individuals responding favorably recorded.

Does this probability experiment represent a binomial experiment?

____No because trials of the experiment are not independent

____No because the probability of success differs from trial to trial

____No because there are more than two mutually exclusive outcomes for each trial

____Yes because the experiment satisfies all the criteria for a binomial experiment.

4. Assume that a procedure yields a binomial distribution with n trials and the probability of success for one trial is p. Use the given values of n and p to find the mean and standard deviation. Also, use the range rule of thumb to find the minimum usual value of of $\mu - 2*\sigma$ and the maximum usual value of $\mu + 2*\sigma$

μ = _____ σ = _____

$\mu - 2*\sigma$ = _____ $\mu + 2*\sigma$ _____

5. Assume that a procedure yields a binomial distribution with a trial repeated n times. Use the binomial probability formula to find the probability of x successes given the probability p of success on a single trial.

- If n = 8, x = 3, and p = 0.015 then compute P(3)

6. The accompanying table describes results from eight offspring peas. The random variable x represents the number of offspring peas with green pods.

a. The probability of getting exactly 7 peas with green pods is ____

b. The probability of getting 7 or more peas with green pods is ____

c. Which probability is relevant for determining whether seven is unusually high number of green pods?

____ The result from part (b)

____ The result from part (a)

d. Is 7 an unusually high number of peas with green pods? Use 0.05 as the threshold for an unusual event.

____ Yes since the appropriate probability is less than 0.05 it is an unusually high number

____ Yes since the appropriate probability is greater than 0.05 it is an unusually high number

____ No since it appropriate probability is less than 0.05 it is not an unusually high number

____ No since the appropriate probability is greater than 0.05 it is not an unusually high number

6. Determine whether the random variable is discrete or continuous in the following examples:

a. Time it takes for a light bulb to burn out ____
b. Number of free-throw attempts before the first shot is made ____
c. Number of points scored during a basketball game ____
d. Number of textbook authors now sitting at a computer ____
e. Amount of snowfall ____

7. Assume that a procedure yields a binomial distribution with a trial repeated n times. Use a binomial probabilities table to find the probability of x successes given the probability p of success on a given trial.

- If n = 6, x = 6, and p = 0.06 then compute P(6) ____

9. A baseball player had 236 hits in a season. In the given probability distribution the random variable x represents the number of hits the player made during a game.

a. Compute and interpret the mean of the random variable x ____

b. Compute the standard deviation of the random variable x ____

b. Which of the following interpretations of the mean is correct:

____ The observed value of the random variable will almost always be equal to the mean of the random variable

____ As the number of trials n increases the mean of all the observations will approach the mean of the random variable

____ The observed value of the random variable will almost always be less than the mean of the random variable

____ As the number of trials n decreases the mean of the observations will approach the mean of the random variable

Statistics in News and Current Events

This section contains news articles that make use of statistics. Please read these articles, look at how statistics are used, and then use your judgement to assess whether statistical information is used accurately and correctly. Be ready to discuss these articles with your professor

1. To Help People Lose Weight Give Them Free Meals and Money, The Wall Street Journal, October 12, 2010
2. To Find Love Match This Valentine's Day Try Love Math, The Wall Street Journal, February 13, 2015
3. At UPS, the Algorithm is the Driver, The Wall Street Journal, February 16, 2015
4. College Majors Figure Big in Earnings, The Wall Street Journal, May 7, 2015
5. Your Cash is No Good Here, The Wall Street Journal, December 28, 2018
6. Google Has Picked an Answer For You, Too Bad It's Often Wrong, The Wall Street Journal, November 16, 2017

This page unintentionally left blank

THE WALL STREET JOURNAL.

https://blogs.wsj.com/health/2010/10/11/how-to-help-people-lose-weight-give-em-free-meals-and-money/

HEALTH BLOG

How To Help People Lose Weight? Give ‘Em Free Meals and Money!

By Katherine Hobson
Oct 11, 2010 12:57 pm ET

Some 68% of the country is either overweight or obese, but there's not much clinical evidence on the best way to lose weight, outside of research on bariatric surgery. (True, "Eat less and exercise more" is great advice if you can manage to follow it, but public-health authorities have repeated that line until they're blue in the face and as a population we've only gotten fatter.)

Research presented over the weekend at the annual meeting of the Obesity Society and published online in the Journal of the American Medical Association suggests a few specific strategies that may help -- namely making a commercial weight-loss program, including packaged foods, free.

In the study, some 442 overweight or obese women were randomly assigned to the Jenny Craig weight-loss program with once-weekly face-to-face counseling sessions, to the same program with weekly phone counseling or to a control group that received "usual care," i.e. publicly available materials on diet and exercise. The two groups using the program also got $25 payments for every 6-month clinic visit to assess weight and other measurements. Jenny Craig, a unit of Nestlé, funded the research.

After two years, the in-person group had lost an average of 16.3 pounds, the phone group lost 13.6 pounds and the control group lost 4.4 pounds.

In the editorial accompanying the study, Rena Wing, a professor of psychiatry and human behavior at Brown University's Warren Alpert Medical School, writes the results suggest that if commercial weight-loss programs such as Jenny Craig were free to those who want to lose weight, "both retention and average weight loss outcomes might be far better than when participants must pay for these programs." The

program costs about $1,600 for 12 weeks, she writes. Compare that to the $19,000 to $29,000 cost of bariatric surgery, which is often reimbursed by insurers.

(The countervailing theory is that people value things more when they pay for them, and that the mere act of ponying up $1,600 for three months of a diet program may make you take it more seriously.)

Wing calls for cost-effectiveness studies, conducted by researchers without any financial ties to the companies involved, that compare different commercial programs with each other. She also notes that more research on weight loss maintenance is needed, since the average weight loss was greater at one year than at the end of the study.

Further reading:

- Hasta La Vista Meridia: Another Diet Drug Bites the Dust
- Bloomberg Says Soda Shouldn't Be Purchased With Food Stamps
- Dieters Foregoing Sleep May Lose Muscle, Not Fat

Share this: http://on.wsj.com/18m1V4S

THE WALL STREET JOURNAL.

https://www.wsj.com/articles/to-find-love-match-this-valentines-day-try-love-math-1423842975

THE NUMBERS

To Find Love Match, Try Love Math (Results Will Vary)

How Lonely Nerds Tailored an Interstellar Formula to Quantify the Astronomical Odds of Finding a Mate

Love Park in downtown Philadelphia **PHOTO:** GETTY IMAGES

By Jo Craven McGinty

Feb. 13, 2015 10:56 a.m. ET

A few years ago, a lonely American in London began to ponder his chances of finding a girlfriend. Surely in a city of 8 million people, plenty of women would find a dark-haired doctoral candidate in economics attractive.

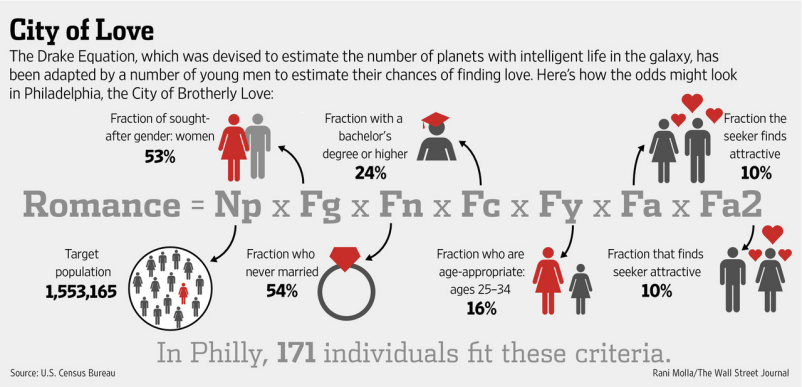
But how many women might be into him? What were his actual odds of finding love?

These are questions more than a few lonely hearts are pondering this Valentine's Day. To try and answer as precisely as possible, Peter Backus, who was 31 at the time, used a mathematical equation to identify the number of women who fit his criteria for romance.

The result, he calculated, was a depressing 26.

Mr. Backus based his calculation on the Drake Equation, a probability formula devised in 1961 by astronomer and physicist Frank Drake to try to solve another great universal mystery: the number of planets in our galaxy that could sustain intelligent life.

In Mr. Backus's romantic version of the Drake Equation, "intelligent alien" equaled "girlfriend."



Here's how the equation works: In a given population—say, London or New York or wherever you happen to live—you apply a series of increasingly restrictive criteria until you end up with a subpopulation that satisfies all the conditions.

The seeker can include any number of criteria—say, the percentage of the population that is college-educated, the number who love "The Hobbit," the fraction who spent their summers at science camp, or the portion who are blond, Republican, Catholic, taller than average or passionate about KenKen math puzzles.

Of course, statistics aren't available for every characteristic, and increasing the number of conditions—being choosier, if you will—quickly shrinks the pool of prospective mates.

In Mr. Backus's case, he wanted to find a single woman around his age in London who was college-educated, attractive and—potential deal-breaker—one with whom he could get along.

The product of the variables produced his result: U.K. population (60,975,000 at the time) x living in London (13%) x women (51%) x single (50%) x between the ages of 24 and 34 (20%) x university educated (26%) x he found attractive (5%) x found him attractive (5%) x compatible (10%).

He obtained most of the figures from the U.K. Office for National Statistics. The three dealing with attractiveness and compatibility were estimates based on his personal experience.

"My friends thought that's a very Peter thing to do," Mr. Backus, who is now married and still living in the U.K., said of his calculation, which he wrote up as a joke and posted online.

But he wasn't alone in his analytical quest for love. Around the same time he deduced his odds for finding a mate, a doctoral student in astrobiology in Seattle wondered how many suitable partners were out there for him. Among his criteria: Must love the Denver Broncos.

Yet another doctoral candidate, studying electrical and computer engineering in Ontario, had the same idea, writing up his assessment in an essay called "Single LGM Seeks Same."

LGM stands for "little green man." Clearly, a tolerance for Martian jokes was a plus.

MORE NUMBERS

- See past Numbers columns
- Check out The Numbers blog

Those three were preceded a decade earlier by a computer-science graduate from Saskatchewan who weighed his chances and, arriving at his mathematical answer, opined that he would never have a girlfriend.

Among his criteria: one standard deviation above the norm for intelligence, and two standard deviations above the norm for beauty.

These lonely young men aren't the only ones to have jotted down the Drake Equation to estimate their chances of finding love.

National Public Radio correspondent David Kestenbaum confided to Ira Glass in a segment of This American Life that he and his Harvard classmates, who were doctoral students in physics at the time, had worked out the equation, too.

And sitcom writer Dave Goetsch introduced the concept in an episode of "The Big Bang Theory" in which a cast of fictional nerds—also a group of physicists and engineers—applied the equation to determine their odds of hooking up with someone on "anything can happen Thursday."

"I thought this is exactly how our characters think," Mr. Goetsch said.

It might be tempting to joke that it is obvious why none of these young men had a romantic partner. In each case—real-life and fictional—each turned not to courtship to solve his problem but to analytics, perhaps with the goal of demonstrating that not having a mate wasn't his fault.

After all, according to their calculations, the odds were against them.

But not so fast! Mr. Backus found and married someone who matched his parameters, and most of the others have also found mates. It seems they have defied the odds—perhaps because at some point they set aside their calculations and concluded love cannot be reduced to a number.

As Raymond Francis, the Ontario student who now, as a postdoctoral scholar, helps support NASA's Mars Science Laboratory, said: "You could spend lot of time trying to nail down those factors. But maybe you should just start looking for someone."

https://www.wsj.com/articles/at-ups-the-algorithm-is-the-driver-1424136536

TECH

At UPS, the Algorithm Is the Driver

Turn right, turn left, turn right: inside Orion, the 10-year effort to squeeze every penny from delivery routes

By Steven Rosenbush and Laura Stevens

Feb. 16, 2015 8:28 p.m. ET

TIMONIUM, Md.—Here's a math problem for you. Each United Parcel Service Inc. UPS **0.16%** ▲ driver makes an average of 120 stops per day. There are 6,689,502,913,449,135,000,000,000,000,000,000,000,000,000, 000,000,000,000,000,000,000, 000,000,000,000,000,000,000,000,000,000,000,000,000, 000,000,000,000,000,000,000,000,000,000,000,000,000, 000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000 alternatives for ordering those stops. Which option is the most efficient, after considering variables such as special delivery times, road regulations, and the existence of private roads that don't appear on a map?

Even if an optimal answer exists, the human mind will never figure it out. And while experts at UPS have been giving the problem their best shot for more than a century, the company is shifting that work over to a computer platform called Orion, which is 10 years and an estimated hundreds of millions of dollars in the making. "Can a human really think of the best way to deliver 120 stops? This is where the algorithm will come in. It will explore paths of doing things you would not, because there are just too many combinations," says Jack Levis, senior director of process management at UPS.

The 1,000 page Orion algorithm is an exercise in heuristics, written by a team of 50 UPS engineers in Timonium, Md. Instead of searching for the optimal, or best possible answer, heuristics is the search for the best answer one can find, the results continually refined over time, based on experience. Orion consists of many components, including a "traveling salesman" algorithm, a familiar tool that calculates the most efficient path between a variety of points, and geographic mapping. What makes Orion unique is the way it puts these elements

together, striving for a balance between an optimum result and consistency, according to Mr. Levis. “Customers and drivers like consistency. Orion has to know when to give up a penny to make the results more stable,” Mr. Levis said.

None of the solutions that Orion spews out are big or dramatic. It is all about saving a dollar or two here and there. But in a network with 55,000 routes in the U.S. alone, that adds up. “In our business, small things mean a lot. If you can re-engineer process, the gains will be greater than you think,” Mr. Levis said.

Such savings matter to UPS, which is struggling with a tighter-margin business and a union workforce that is compensated at the high end of the industry scale. Its challenges are unique. Rival FedEx Corp. uses an independent contractor model for its ground network, so it isn’t ultimately responsible for miles driven to most of its residential stops.

A Changing Business

E-commerce has shifted more and more of UPS’s delivery stops to residences, and those packages are expected to make up half of all deliveries by 2018. It is a radical change from 15 years ago, when drivers would drop off several packages at a retailer. Now, they make scattered stops to drop off one package at houses in a neighborhood, driving further and taking up more time.

On Nov. 13, UPS CEO David Abney said he expected Orion to save the company $300 million to $400 million a year, once it is fully implemented in 2017. The more than 40% of the company’s 55,000 U.S. routes already using the software at that time had been reduced by an average of between seven and 8 miles, the company said. The company can save $50 million a year by reducing by one mile the average aggregated daily travel of its drivers. Those savings are critical as UPS tries to boost earnings growth, which has been in the 5% range in recent years and dipped in 2014, as low-margin deliveries related to e-commerce become more prevalent and the company scrambles to figure out how to manage its holiday season.

While Mr. Abney cautioned that at least some of Orion’s gains would be offset by rising costs related to delivery of its customers’ e-commerce orders, he is targeting per-share earnings growth of more than 50% over the next five years. The company lowered its 2015 outlook earlier this month.

UPS won’t say how much money it has invested in Orion. But management and information technology expert Thomas H. Davenport, a distinguished professor at Babson College near Boston, believes Orion is the largest deployment of operations research, and that UPS spent $200 million to $300 million to develop it, excluding many years of investments in underlying driver technology and communications infrastructure.

How Orion works

A driver—in this case, let's use the example of Tim Ahn, who has been a full-time driver for 20 years, currently with a route in Gettsyburg, Pa.—would use his UPS tablet, known within the company as a delivery information acquisition device, or DIAD, to punch in at the beginning of his shift, as he does now. The DIAD would show him two possible ways to make his deliveries, one using Orion, and one using the current combination of work rules, procedures and analytic tools that are used to establish the order of package deliveries. He can choose to work in either way, but if he decides not to use Orion, he will be asked to explain the decision.

Orion already has been at work for hours, though. It may have reordered Mr. Ahn's schedule of stops for the day hundreds of times, as packages were added to the list assembled before he arrives at work, and as customers used the company's My Choice self-service platform to change the time or location of their deliveries. UPS says My Choice membership has grown steadily since its launch in 2011 to 12.9 million today.

At one point, Mr. Ahn was scheduled to start his route at 8:45 a.m., making 125 deliveries and traveling 117.85 miles during the day. But now Customer 1 wants a package delivered between 11 a.m. and 1 p.m. That stop was originally scheduled by Orion for 1:25, so Orion has to recalculate. It considers up to 200,000 of the best options before settling on one. The package will now be delivered by 12:30 p.m., adding 1.39 miles to the day's route, at a cost to UPS of $1.99. It takes Orion and the network about eight seconds to return an answer.

Now, Customer 2 specifies that a package that Orion originally scheduled for delivery by UPS at 3:51 p.m. must take place between 4:30 p.m. and 6:30 p.m. Orion considers a range of options before settling on a delivery order that arranges the delivery for 4:46 p.m., adding 1.64 additional miles and $2.77 in cost.

Orion is a useful tool, according to Mr. Ahn. "Orion had me do things in the morning I would not think of doing, and it saved me miles later in the day," he said.

Rough Patches

The deployment of Orion isn't always so smooth, though. That is where Mr. Levis comes in. As project manager, he is responsible for getting people and machines to work together. During the earlier stages of writing the Orion algorithm, it was Orion that had to learn to accommodate people.

Jack Levis is the leader of UPS's Orion project, which has been some 10 years in the making. PHOTO: MATT ROTH FOR THE WALL STREET JOURNAL

"The project was nearly killed in 2007, because it kept spitting out answers that we couldn't implement," Mr. Levis recalls. The earliest versions of Orion focused on getting the best mathematical results, with insufficient regard for the interests of the driver or the customer, who value some level of routine. For example, regular business customers who receive packages on a daily basis don't want UPS to show up at 10 a.m. one day, and 5 p.m. the next. And a customer who is expecting a shipment of frozen food needs delivery as soon as possible, even if efficiency demands that someone gets priority.

To get the project back on track, UPS chief scientist Ranga Nuggehalli turned to Bob Santilli, a senior project manager, asking him to describe a perfect route. Several weeks later, Mr. Santilli came back with the results of his effort, which produced a model plan of stops for drivers on a route in Lancaster, Pa. The engineering team extracted proprietary rules from the Santilli route and built them into Orion.

"By April or May of 2007, he had the first working version of Orion, which balanced consistency and optimality. It had to do with keeping the driver in a path. The route should flow. That is what we learned. That is what brings consistency. Orion can make exceptions to the flow, but it has to do so in an intelligent manner and it can't make an unlimited number of exceptions," Mr. Levis said.

55,000 Routes

The process of balancing Orion's logic with the real-world experience of drivers is built into the rollout of the project. A team of 700 trainers is working its way through all 55,000 U.S. routes, deploying Orion to one UPS facility at a time, a process expected to be more than 70% complete by the end of the year.

It takes about six days to train a driver. The first day of training is spent fixing maps, as the trainers pore over satellite images and talk to drivers about minute details of their routes. On the third day, the trainers ride the route themselves in a rental car. On the fourth and fifth day, the trainers ride with the driver, and try to figure out what Orion is getting wrong about the route. More revisions are made on the fifth day, and a final ride-along occurs on day six.

BUSINESS & TECH.

- Introducing a New Name: Business & Tech.
- Challenge of Apple Watch: Defining Its Purpose
- Landmark Rules for Commercial Drones
- Keywords: You Aren't a Human, You're a Data Point

Driver reaction to Orion is mixed. The experience can be frustrating for some who might not want to give up a degree of autonomy, or who might not follow Orion's logic. For example, some drivers don't understand why it makes sense to deliver a package in one neighborhood in the morning, and come back to the same area later in the day for another delivery. But Orion often can see a payoff, measured in small amounts of time and money that the average person might not see.

United Parcel Service driver Marty Thompson steps off a truck while making a delivery in Cumming, Ga., in June. PHOTO: ASSOCIATED PRESS

Logical or Illogical?

One driver, who declined to speak for attribution, said he has been on Orion since mid-2014 and dislikes it, because it strikes him as illogical. He said that while a colleague who drives a rural route saves more than 20 miles a day using Orion, the program actually added miles to his urban routes when it reduced the total number of routes and combined them. He says the

program calculates routes with more left turns and assumes he'll be backing up—two things UPS drivers are taught to avoid to keep safe. And he doesn't like it when Orion tells a driver to deliver to a neighborhood but skip some houses, leaving some stops in the area for another driver.

A second driver who started on Orion this year echoed similar concerns.

A UPS spokeswoman said that drivers are also supposed to use their own judgment in following Orion, and that the program doesn't direct them to violate safety rules.

For example, drivers could refrain from using Orion if there is a traffic event that the system can't factor. But the company maintains that a driver together with Orion is better than each alone.

Like it or not, more automation is coming to UPS.

"Orion...is not an endgame; it is part of a platform," Mr. Abney, UPS's CEO, said. "[T]hese initiatives, along with others, will reduce our delivery costs and provide economic value to our customers and our shareholders."

UPS engineers are already enhancing Orion so it will update delivery schedules while drivers are on the road, useful in a situation in which a driver might abandon Orion's instructions because of an unexpected road closure due to an accident, but want to resume using Orion later in the day. Upcoming versions also will include turn-by-turn driving instructions—not yet part of the system.

At some point, Orion may coincide with the rise of driverless vehicles. While true self-driving cars won't be on the road any time soon, the idea of connecting a few driverless trucks in a platoon with one driver in a vehicle at the front isn't far-fetched, according to Mr. Davenport. "What must be scary is that there will be automated vehicles at some point, although my guess is that it will not happen any time soon," he says. "The driver will have less and less to do."

Write to Steven Rosenbush at steven.rosenbush@wsj.com and Laura Stevens at laura.stevens@wsj.com

THE WALL STREET JOURNAL.

https://www.wsj.com/articles/college-majors-figure-big-in-earnings-1430971261

EDUCATION

College Majors Figure Big in Earnings

Study finds some study areas pay more than others, with engineering earnings triple those for education

By Melissa Korn

Updated May 7, 2015 11:41 a.m. ET

Want to make a good living? Go to college. Just be careful what you major in.

On average, college graduates earn about $1 million more in their lifetimes than do adults who only completed high school. But long-term earnings prospects vary widely by subject, and the income differentials across certain majors dwarf those between graduates and non-graduates, according to a new report from Georgetown University's Center on Education and the Workforce based on an analysis of Census Bureau data.

For example, students who complete undergraduate degrees in petroleum engineering earn a median $4.8 million throughout their careers (or $136,000 a year)—more than triple the $1.4 million in median earnings (or $39,000 a year) for someone who majored in early-childhood education, the report says.

The findings, which include detailed analyses of earnings for graduates in 137 specific majors, add fuel to an already heated debate over the value of a college education amid skyrocketing student-loan debt and a still-tepid job market for fresh graduates.

Think tanks, states and the federal government have scrambled to fill the information void with a variety of tools to help families try to measure the return on their college investments. The Brookings Institution last week ranked schools based on their "value-added," or the difference between mid-career salaries of alumni and estimated salaries of comparable students from

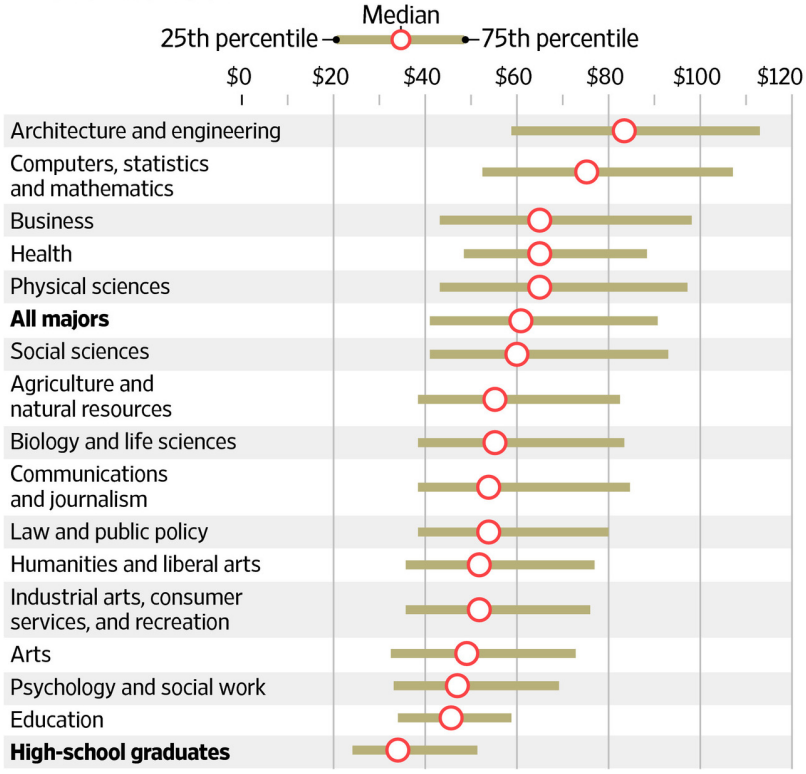
Amy Lawson, a fifth-grade teacher at Silver Lake Elementary School in Middletown, Del. in 2013, is likely to earn less than someone who studied petroleum engineering. **PHOTO:** STEVE RUARK/ASSOCIATED PRESS

other schools. The California Institute of Technology topped the list of four-year schools with the most value added in terms of mid-career earnings.

Income Inequality

Going to college pays off, but by how much depends greatly on the area of study.

Annual wages of college graduates by major over a career (ages 25–59)
In thousands of dollars



Source: Georgetown University THE WALL STREET JOURNAL.

Georgetown published a report in February looking at the wage premiums for recent college graduates compared with high-school graduates, along with the unemployment rates for various majors. Graduates of architecture programs may have higher salaries than teachers, as the latest paper shows, but the February report noted that they're also likely to see unemployment rates twice those of education majors.

The latest Georgetown report shows that sweeping statements about college graduates' earnings—whether based on first-year outcomes for an entire school or even averages within fields of study—say little about prospects for individual graduates.

ON TOTAL RETURN »

- Financial Advice for New College Grads

"The surprises are in the details," said Anthony Carnevale, director of the Georgetown center. For example, just choosing a major in science, technology, engineering or mathematics, known as the STEM fields, doesn't secure a hefty paycheck. Mr. Carnevale's team found that biology majors have median annual wages of $56,000 over their careers from age 25 to 59, or about one-third less than physicists.

Yet once biologists finish graduate programs—and more than half of them do—their median annual earnings jump to $96,000, roughly on par with physicists who have advanced degrees.

There are also wide ranges in salaries for specific majors. The top 25% of earners who majored in finance can expect annual earnings of more than $100,000, while the bottom quartile may bring in just about $50,000 a year.

College graduation is a good thing, but how it pays off in lifetime earnings depends to a large degree on a student's major subject, a Georgetown University study finds. PHOTO: EMILY VARISCO/ASSOCIATED PRESS

"When people see a median, they think it's destiny," Mr. Carnevale said. "It's not. There are people above and people below."

Similarly, the Brookings Institution's Hamilton Project found last fall that lifetime earnings for economics majors at the 90th percentile are nearly triple those at the 10th, reflecting the range of destinations for such experts in government and the private sector.

Some states are trying to encourage students to pursue certain degree programs and steer clear of ones that may lead to lower-paying jobs or higher unemployment rates.

North Carolina Gov. Pat McCrory has stated that he would like to allocate higher education funding based on job prospects for certain programs, and proposed in next year's budget to upgrade STEM-department technology at state universities to help attract more students.

Florida currently weighs graduate outcomes in determining school spending under its performance-based funding model. “I love all degrees,” Gov. Rick Scott said in an interview with The Journal Wednesday. “But I like the degrees where kids get jobs.”

Melissa Kearney, an economics professor at the University of Maryland and director of the Hamilton Project, said salaries should factor into, but not necessarily dominate, decisions on majors.

> **Yes, engineering pays very well. But that’s not going to serve someone who really struggles with quantitative skills.**
>
> —Melissa Kearney, director, Brookings Institution’s Hamilton Project

“People have to think very carefully about what they’re good at,” she said. “Yes, engineering pays very well. But that’s not going to serve someone who really struggles with quantitative skills.”

Low earnings prospects are a concern for Anne Davis, a junior studying sociology and environmental policy at the College of William & Mary in Williamsburg, Va. But she is still pursuing a career in grassroots organizing for environmental groups.

Ms. Davis, 21 years old, hopes to land a full-time job working as a community organizer at an environmental nonprofit, where she expects to earn $35,000 to $40,000.

“I’ve found my calling,” she said. “I know the money’s not lucrative, but I couldn’t see myself doing anything else.”

Write to Melissa Korn at melissa.korn@wsj.com

WSJ ›

Annotations

Your Cash Is No Good Here. Literally.

KATHERINE BINDLEY DECEMBER 28, 2018

Sam Russell/PA Wire/Zuma Press

Sam Schreiber was mid-shampoo at a Drybar blow-dry salon in Los Angeles when someone from the front desk approached her stylist with an emergency: a woman was trying to pay for her blow-out with cash.

"There was this beat of silence," says Ms. Schreiber, 33 years old. "She literally brought $40."

More and more businesses like Drybar don't want your money—the paper kind at least. It's making things awkward for those who come ill prepared.

After all, you can't give back a hairdo, an already dressed salad or the two beers you already drank.

The salad chain Sweetgreen has stopped accepting cash in nearly all its locations. Most Dig Inns—which serve locally sourced, healthy fast food—won't take your bills either. Starbucks went cashless at a Seattle location in January, and at some pubs in the U.K., you can no longer get a pint with pound notes. The practice of not accepting cash has become popular enough to catch the attention of American lawmakers.

Sam Schreiber helped out a fellow customer who only had cash at Drybar.

Photo: David Baggelaar

Ms. Schreiber was tempted to wait and see how the Drybar employees would handle the situation with the customer, who had no credit or debit

card with her; instead, she intervened from the shampoo bowl. "I said, 'I can just pay for her and she can give me cash or Venmo me,' " she says.

A few moments later, one of the employees came back to hand her the $40 and expressed thanks on behalf of the stranger. The staff also offered her a second mimosa, which are free to customers. "I kind of wanted to be, like, I should get a free updo or something," she says. "I basically was…the bank for them." Drybar declined to comment.

Despite the popularity of debit- and credit-card transactions, plenty of people do still pay for things with actual money. Cash represented 30% of all transactions and 55% of those under $10, according to a Federal Reserve survey of 2,800 people conducted in October 2017.

Jaclyn Benton, 30, visited a Sweetgreen near her office in Reston, Va., last summer with $20 cash, but no credit or debit card because she had forgotten her wallet at home. When her order was ready and she went to pay, the cashier explained that the restaurant doesn't take bills.

"It's almost like when your credit card gets declined for silly reasons," says Ms. Benton, who works as an event planner. "It makes you feel like you can't afford it even though I had the money right there."

Ms. Benton has no plans to go back: "It feels very elitist," she says.

A Sweetgreen spokeswoman said its decision makes its team members safer amid the risk of robbery and improves the cleanliness and efficiency of the restaurants.

The Boot pub in Freston, England, posts a sign to let customers know it doesn't take cash. Owner Mike Keen used his own card to cover a customer who only had notes. Photo: Sam Russell/Zuma Press

Brooke Linbourne, 51, ended up strolling with a credit-card-bearing stranger in search of a place to break a $20, after an "awkward" encounter at a Dig Inn in Manhattan. She said she slowly became aware while waiting in line to pay that all the other customers were using cards. At the register, the woman behind her in line helped her out.

"We went into one of those stores where they sell Lotto tickets and I got change and I gave her the money," says Ms. Linbourne, who lives in Hasbrouck Heights, N.J., and works for a construction-management company. "I was so embarrassed."

A Dig Inn spokeswoman referred to a company blog post about why it has cashless locations. It says it instituted the policy in locations where cash made up 8% or less of its transactions, and says it makes for a faster experience for customers and for employees, who don't have to count cash or make runs to the bank.

Although U.S. bills feature the words, "This note is legal tender for all debts, public and private," there's no federal law that says businesses have to accept cash, according to the Federal Reserve's website.

The salad chain Sweetgreen has stopped accepting cash in nearly all its locations. Photo: Richard B. Levine/Newscom/Zuma Press

Massachusetts is the only state that currently requires retailers to accept cash. Some New Jersey legislators are working to make their state next. New York City Councilman Ritchie Torres of the Bronx recently proposed legislation that would prohibit retailers and restaurants from refusing cash, and city council members in Washington, D.C., and Philadelphia have proposed similar legislation.

"I refuse to patronize businesses that reject cash payments, even though I primarily use debit or credit," says Councilman Torres. He says not accepting cash is discriminatory against the undocumented, people without bank accounts and credit cards, and those who wish to have their transactions be more private. "It's a humiliating situation."

Another demographic that often only has cash? Minors. Connie Young, who lives in Walnut Creek, Calif., says that in February, her 17-year-old son got excited when he learned a book he wanted was in stock at the local Amazon Books.

But her son returned home empty handed. When he told her the store didn't take cash, she assumed there must have been a power outage and that the register was down, before he explained it was the policy. "I laughed. I was, like, you're kidding," says Ms. Young, 57. "I was just stunned."

The Boot pub in Freston, England, which opened about six months ago, decided to go cashless because it saves employees time and lowers insurance premiums, says its owner, Mike Keen.

Connie Young's 17-year-old son couldn't get a book he wanted at an Amazon Books because it didn't take cash.Photo: Nancy Kaszerman/Zuma

Press

Once in a while someone isn't aware of the policy. In September, a customer drank two pints, asked for his bill and handed over a £10 note, which Mr. Keen explained he couldn't accept. He says the patron was sitting near a blackboard that read, "First cashless bar in England."

"He's extremely embarrassed and patting his pockets to see if he's suddenly sprouted a card," says Mr. Keen. "Funnily enough, he hasn't in the five minutes that we've been talking."

Mr. Keen offered to put the customer's bill on his own credit card and pocket the cash. "I'm not a bank myself but if it came down to accepting the money or not, then that's the only way," he says.

Yassim Alsalman, a rapper and multimedia artist from Canada whose stage name is Narcy, doesn't bring a credit card while traveling to hold himself to a budget. It worked to his advantage during a visit to New York. He was offered a free smoothie in Brooklyn, followed by a free salad from Sweetgreen later that afternoon. Both establishments handed him the goods after he produced a $20 bill and said he only had cash. "I had a full day of lunch," he says.

https://www.wsj.com/articles/googles-featured-answers-aim-to-distill-truthbut-often-get-it-wrong-1510847867

TECH

Google Has Picked an Answer for You—Too Bad It's Often Wrong

Going beyond search, the internet giant is promoting a single result over all others, and many are contentious, improbable or laughably incorrect

By Jack Nicas

Nov. 16, 2017 10:58 a.m. ET

Google became the world's go-to source of information by ranking billions of links from millions of sources. Now, for many queries, the internet giant is presenting itself as the authority on truth by promoting a single search result as the answer.

To the question "Does money buy happiness?" Google recently highlighted a result that stated: "There is enough scientific research to prove" it.

"Who are the worst CEOs of all time?" Google answered with the names and photos of 11 chief executives, including Gordon Bethune of Continental Airlines and Robert Nardelli of Home Depot Inc.

Sometimes, Google's response depends on how the question is asked. For "Should abortion be legal?" Google cited a South African news site saying, "It is not the place of government to legislate against woman's choices."

When asked, "Should abortion be illegal?" it promoted an answer from obscure clickbait site listland.com stating, "Abortion is murder."

The promoted answers, called featured snippets, are outlined in boxes above other results and presented in larger type, often with images. Google's voice assistant sometimes reads them aloud. They give Google's secret algorithms even greater power to shape public opinion, given that surveys show people consider search engines their most-trusted source of information, over traditional media or social media.

Google typically lists the source below the answer—or credits the source first when reading an answer aloud—but not always. The worst-CEOs list was unsourced. "That's the dumbest bunch of shit I've ever seen," Mr. Bethune said in an interview. Mr. Nardelli declined to comment.

Also unsourced was an inaccurate answer that said former President Barack Obama and Rep. Peter King (R., NY) are Muslim members of Congress.

Google's featured answers are feeding a raging global debate about the ability of Silicon Valley companies to influence society. Google and other internet giants are under intensifying scrutiny over the power of their products and their vulnerability to bias or manipulation.

Facebook Inc. was criticized for enabling the spread of false news reports during the 2016 presidential election, and Google has been called out for promoting discredited conspiracy theories, including about recent mass shootings in Nevada and Texas. Executives from Google, Facebook and Twitter Inc. were called before Congress in recent weeks to testify about Russia-backed accounts that used their platforms to sow misinformation. The companies say they are addressing the issues.

Google spokeswoman Susan Cadrecha said the company's goal isn't to do the thinking for users but "to help you find relevant information quickly and easily." She added, "We encourage users to understand the full context by clicking through to the source."

Featured snippets are "generated algorithmically and [are] a reflection of what people are searching for and what's available on the web," the company said in an April blog post. "This can sometimes lead to results that are unexpected, inaccurate or offensive."

Google, a unit of Alphabet Inc., handles almost all internet searches. Featured snippets appear on about 40% of results for searches formed as questions, according to a July study completed for The Wall Street Journal by search-data firm SEMrush.

An algorithm chooses featured snippets from websites in part by how closely they appear to satisfy a user's question, factoring in Google's measure of a source's authority and its ranking in the search results.

By answering questions directly, Google aims to make the search engine more appealing to users and the advertisers that chase them. The answers' real estate is so attractive that there is

a budding marketing industry around tailoring content so it becomes a featured snippet. (There is even a featured snippet for "how to get a featured snippet.")

Digital-marketing firm Stone Temple Consulting, which tracked nearly 1.5 million searches, found that as Google expanded the use of featured snippets, it has relied more often on less authoritative sources, such as purveyors of top-10 lists and gossipy clickbait.

Those issues have spurred an internal debate on Google's search team over how much they should meddle with the featured answers, which the group believes have a greater weight with users than typical search results, according to a former manager on the team.

University of North Carolina professor Zeynep Tufekci, who studies technology's effect on society, said Google shouldn't put its trusted seal of approval on answers it isn't certain are accurate. "For them to wield their algorithm like this is very worrisome," she said. "This is how people learn about the world."

Gummi Hafsteinsson, who oversees Google's virtual assistant, said in an interview that teams of Google employees try to weed out inaccurate answers, but that the answers overall help inform because they are almost always right. "The kinds of things we can answer are unbelievable," he said. No one at Google writes an answer for "how to remove a red stain from a carpet," he said, but Google's algorithm finds a solution on the web and serves it to users.

"I think the benefit is tremendously big," he said. "It's always a balancing act in terms of quality."

A study this year by Stone Temple, a prominent analyst of the industry, showed Google's search engine answered 74.3% of 5,000 questions, and on those answers it had a 97.4% accuracy rate. Both percentages are higher than services from Amazon.com Inc., Apple Inc. and Microsoft Corp.

Yet since Google handles trillions of queries a year, even a 2.6% error rate suggests Google serves billions of answers a year that are incomplete, irrelevant or wrong.

An Amazon spokeswoman said it answers most questions with information from trusted third parties, and that it is careful with answers on sensitive topics. An Apple spokeswoman said it generally only answers questions for which it has clear factual answers. Microsoft said it aims to offer answers on its Bing search engine and Cortana virtual assistant that are "relevant, balanced and trustworthy."

Google launched in 1998 and quickly attracted users. Its simple lists of blue links distilled the internet's immense content, ranked by an algorithm based on how often websites cited each other.

In 2012, Google began answering basic factual queries with so-called knowledge cards, drawn from an internal encyclopedia called the knowledge graph that has more than a billion entries based on sources including Wikipedia and the Central Intelligence Agency's World Factbook. For questions with clear answers, such as, "How tall is Shaq?" the knowledge graph proved reliable in helping users quickly find information. But it could handle only a small fraction of questions.

At the time, the company was developing Google Glass, a wearable computer resembling eyeglasses that placed a tiny screen in the user's peripheral vision. The screen couldn't display a webpage or scroll through lists, but it could answer questions.

About five engineers from the several thousand employees on the search team began developing a question-answer system that pulled answers from a wide range of internet sources, according to the former search manager.

Google Glass was a flop, but the search team saw the value of the answer system and made it a focus, the person said.

That proved prescient. Today, Google is locked in a race with four other U.S. tech giants—Apple, Microsoft, Amazon and Facebook—to win users with more intelligent services, including virtual assistants, that they hope will make them more central to users' lives, and open new opportunities to sell ads and products.

The assistants are often built into devices with smaller screens or no screens at all, including smartphones, watches and voice-controlled speakers like the Amazon Echo and Google Home, where long lists of ranked sources are impractical.

"Searching on a mobile device is very different from a desktop computer. Speed and simplicity really matter," Alphabet Chairman Eric Schmidt said in 2014, the year Google launched featured snippets. "It's why the best answer is quite literally the answer."

With many tech companies betting virtual assistants are the future of computing, question-answer systems are likely to become more common, and wield more influence, in the years ahead.

When Google debuted featured snippets, they appeared on roughly one of every 1,000 searches, the former manager said. The team tweaked the algorithm to "squeeze a little bit more out of it" and increased the number of snippets by roughly 5% to 10% each month, the person said. Subjects expanded to include health, law, business, politics and religion.

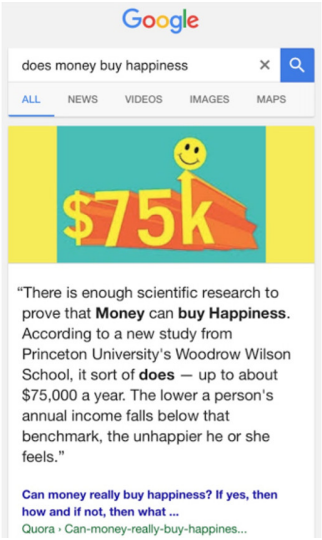
Stone Temple's data indicate Google over time has pulled more answers from less reliable sources. The firm judged the sources using third-party ratings that approximate how Google measures a site's authority.

Answers generated from sources with a 90 or higher rating, such as Wikipedia or the Journal, dropped to less than 40% of featured snippets in February from more than 60% in July 2015. Answers from sites rated below 40, including blogs and clickbait sites, rose to 15% from 6.5% over the period.

Sites with low rankings can generate unreliable answers. To the query "Why are Komodo dragons endangered?" the featured answer was volcanoes, fire and tourism. The source? A Canadian elementary school student's report posted online. Komodo dragons aren't endangered.

An Answer for Everything

Google is increasingly elevating a single search result for many questions and giving it the appearance of a definitive answer. Sometimes the information is controversial, misleading or just bizarre.



Google promoted this answer to an abstract question in a 'featured snippet.' GOOGLE

1 of 6

Because Google's algorithm seeks answers that closely match users' questions, its responses often reflect how a question is framed. That can lead to different answers to similar questions, and contribute to confirming biases.

A recent search for "Is milk good for you?" yielded an answer from a health site saying, "Milk can be good for the bones because it provides vitamin D and calcium."

A separate search for "Is milk bad for you" featured an answer saying, "Calcium from animal milk is not absorbed as well as that from plant-based sources, and it can be accompanied by a number of dangerous health problems." That came from People for the Ethical Treatment of Animals, a group that advocates against any consumption of animal products.

Jackson Miller, owner of a Nashville, Tenn., resale shop, said Google provided the incorrect dates of Tennessee's tax-free weekend this summer, one of the busiest weekends for retailers.

Mr. Miller said he suspects some customers were fooled, because everyone trusts Google. "You treat it as a primary source," he said. "It's like, 'Google says it, so it must be true.' "

Google is constantly changing its search results, so such results only appear some of the time. In the April blog post, Google said it returned "offensive or clearly misleading content" for one in every 400 queries. It said it would improve its algorithm and make it easier for users and employees to flag problem results.

The post came after several inaccurate and unsavory featured snippets got attention on social media, including results saying that several past U.S. presidents were Ku Klux Klan members and that women are evil.

> Ms. Cadrecha of Google told the Journal this week that the company recently changed its algorithm to limit featured snippets on sensitive topics, such as religion and politics.

Mr. Hafsteinsson of Google said the system is designed to avoid unanswerable queries, but while a subjective question seems obvious to humans, "it might not be to the algorithms."

Meanwhile, Google has expanded another element—"People also ask" boxes—that serves up answers to questions similar to a given search. The product appears to rely on the same algorithm as featured snippets and can push misleading information on topics users weren't even searching for.

To a search for, "Are people born evil," a box suggested the question, "Can a person be born homosexual?" Google, citing a website procon.org that presents differing opinions on controversial topics, answered that while "many ex-gays" say they were born gay, "the reality is that no scientific evidence has established a genetic cause for homosexuality."

In February, Google included "People also ask" boxes in 16.3% of its search results, up from 1.4% a year earlier, according to Stone Temple data.

Promoting such answers suggests "they've given it their stamp of approval, to say this is the one versus these are the 10," said Pete Meyers, an analyst who studies Google results for the

marketing-analytics firm Moz Inc. "People generally trust Google, but now these answers aren't coming from a trusted source."

Write to Jack Nicas at jack.nicas@wsj.com

Appeared in the November 17, 2017, print edition as 'Google Has Chosen an Answer For You—It's Often Wrong.'