

Introduction to Statistics for Business and Economics Workbook 1 (Version 1.0)

Peter Dalley

© Peter Dalley, 2015

Table of Contents

Module 1: Types of Data.....	4
Module 2: Graphical Summaries	10
Module 3: Numerical Summaries	22
Module 4: Introduction to Probability.....	38
Module 5: Discrete Probability Distributions.....	53
Module 6: Continuous Probability Distributions.....	68
Module 7: Sampling and Sampling Distributions	79
Module 8: Estimating Intervals	90

A Note to Instructors

I hope you find this workbook useful, I just want to point out three key features:

This book is totally free to you and your students. Feel free to copy it or post it to your course website and feel free to share it with colleagues.

Although I am widely distributing a PDF file, I have gone to great effort to make a fully editable Word version of this document. Please contact me if you'd like to have a copy of the Word version. You can edit any of these problems to better fit in your class or simply copy and paste an entire problem into an assignment or test, with the attribution "Source: statisticsworkbook.com", or "Adapted from: statisticsworkbook.com".

Every problem in this workbook has a video walkthrough available at <http://statisticsworkbook.com>. I suspect the true value in this book lies in the video walkthroughs, as it will be useful for homework and particularly useful for "flipping the classroom".

Please let me know if you would like to see additional question-types or topics included in the future. I intend to add to this book frequently based on your input. Also, any feedback you can provide (particularly student feedback) would be greatly appreciated.

Please note, you do not have my permission to use this for a commercial purpose, nor do you have permission to recreate the videos found at <http://statisticsworkbook.com>. Send me an email if you have any questions about use or attribution.

Thanks for checking out this workbook, and I hope you'll have a look at the companion website: <http://statisticsworkbook.com>!

Peter Dalley

profdalley@gmail.com

Module 1: Types of Data

1-1A Types of data and measurement scales

The following table provides a number of pieces of information on a variety of different vehicles. Respond to the questions below.

Model	Engine size	Cylinders	Transmission	Number of gears	Fuel	Vehicle Class	City MPG	Hwy MPG	Model Year
ACURA RDX	3.5	6	Automatic	6	Gasoline	small SUV	19	28	2010
HYUNDAI Sonata	1.6	4	Manual	7	Gasoline	large car	28	38	2015
FORD Fiesta	1.6	4	Manual	5	Gasoline	small car	28	36	2006
DODGE Challenger	6.4	8	Automatic	8	Gasoline	midsize car	15	25	2010
BMW X5 xDrive35i	3	6	Automatic	8	Gasoline	standard SUV	18	24	2013
HONDA Accord	3.5	6	Automatic	6	Gasoline	midsize car	21	32	2014
LAND ROVER Range Rover	3	6	Automatic	8	Diesel	standard SUV	22	29	2008
BENTLEY Mulsanne	6.8	8	Automatic	8	Gasoline	midsize car	11	18	2012
MAZDA CX-5	2.5	4	Automatic	6	Gasoline	small SUV	24	30	2013
PORSCHE Cayman GTS	3.4	6	Manual	6	Gasoline	small car	19	26	2015

Source: <https://www.fueleconomy.gov>

- How many elements are in the dataset?
- How many variables are in the dataset? What are they?
- What type of measurement scale is used for each variable?
- Which variables are quantitative and which are categorical?

1-1B Type of data and measurement scale

The following table provides some information on universities I couldn't get into.

Title	Country	No. of FTEs	No. of students FTEs per faculty	International Students	Female:Male
California Institute of Technology	USA	2,243	6.9	27%	33:67
University of Oxford	United Kingdom	19,919	11.6	34%	46:54
Stanford University	USA	15,596	7.8	22%	42:58
University of Cambridge	United Kingdom	18,812	11.8	34%	46:54
MIT	USA	11,074	9	33%	37:63
Harvard University	USA	20,152	8.9	25%	n/a
Princeton University	USA	7,929	8.4	27%	45:55
Imperial College London	United Kingdom	15,060	11.7	51%	37:63
ETH Zurich - Swiss Federal Institute of Technology	Switzerland	18,178	14.7	37%	31:69
University of Chicago	USA	14,221	6.9	21%	42:58

- How many elements are in this dataset?
- How many variables are in this dataset? What are they?
- What measurement scale is used for each of the variables?

1-1C Type of data and measurement scale

The following questions were asked on a student drug use survey. State whether the information gathered from them is categorical or quantitative and indicate the appropriate measurement scale for each.

- a) Are you Male or Female?
- b) What grade are you in?
- c) What is your grade point average? (Provide numerical average and corresponding letter grade)
- d) How old were you when you first drank alcohol?
- e) On a scale of 1-5 (1 – Daily, 2 – Weekly, 3 – Monthly, 4 – Annually 5 – Only once in your life):
 - a. rank your frequency of use for alcohol.

 - b. rank your frequency of use for marijuana.

1-2A Time Series and Cross Sectional Data

The following table provides information on National Lunch program participants and populations of three states.

State	National School Lunch Program participants (all values are in millions)			
	2011	2012	2013	2014
California	3.28	3.33	3.29	3.24
Florida	1.65	1.65	1.64	1.66
New York	1.82	1.80	1.72	1.70
	Population			
California	37.69	38.00	38.33	38.80
Florida	19.06	19.32	19.55	19.89
New York	19.47	19.58	19.65	19.75

- What are the variables of interest? What are the elements?
- Is the data categorical or quantitative?
- Is the data cross-sectional or time series?

1-2B Time Series and Cross Sectional Data

The following table shows income per person (Fixed 2000 \$US).

	Gross Domestic Product per Person (2000, \$US)						
	2005	2006	2007	2008	2009	2010	2011
Canada	25437.65	25943.74	26229.74	26101.76	25069.87	25575.22	25933.29
China	1464.108	1640.862	1864.103	2032.615	2208.404	2426.332	2639.542
France	22734.23	23133.35	23516.22	23366.4	22508.76	22758.15	23016.85
Italy	19781.96	20102.23	20291.23	19903.46	18697.22	18944.41	18937.25
Russia	2442.963	2654.2	2888.847	3043.666	2806.415	2928.005	3054.728
United States	37718.01	38349.4	38710.89	38208.76	36539.23	37329.62	37691.03

- What are the variables of interest? What are the elements?
- Are the data cross-sectional or time series?

Module 2: Graphical Summaries

2-1A Summarizing Categorical Variables – Frequency Distributions and Pie Charts

Students of a first year course in Statistics for Business and Economics were asked to rate their satisfaction of the course on the following scale:

1- Very unpleasant 2- Unpleasant 3- Not bad 4- Interesting 5- Very interesting.

The responses of 10 students are as follows:

Unpleasant	Interesting	Not bad	Very unpleasant	Interesting
Interesting	Not bad	Very interesting	Interesting	Unpleasant

a. Fill in the following table:

	Frequency	Relative Frequency	Percent Frequency
Very Unpleasant			
Unpleasant			
Not Bad			
Interesting			
Very Interesting			

- Produce a bar graph to illustrate the frequency of each response.
- Produce a bar graph to illustrate the relative frequencies.
- Produce a pie chart to illustrate the percent frequencies.

2-1B Summarizing Categorical Variables – Frequency Distributions and Pie Charts

A group of students in a first year course in Statistics for Business and Economics received the following letter grades:

D F C B D
A B C C F
D A B C B
A C B F D

- a. Fill in the following table:

	Frequency	Relative Frequency	Percent Frequency
A			
B			
C			
D			
F			

- b. Produce a bar graph to illustrate the frequency of each response.
c. Produce a bar graph to illustrate the relative frequencies.
d. Produce a pie chart to illustrate the percent frequencies.

2-2A Summarizing Quantitative Variables – Bar graphs

The following table provides information on carbon dioxide emissions in 2013 for North American countries.

	CO2 emissions (thousands of tonnes)	Emissions per capita (tonnes)
Canada	565,000	15.9
Mexico	456,000	3.7
United States	5,335,000	16.5
North America	6,356,000	13.56

- a) Construct a bar graph showing CO2 emissions for Canada, Mexico and the United States.
- b) Construct a bar graph showing per capita CO2 emissions for Canada, Mexico and the United States.

2-2B Summarizing Quantitative Variables – Bar graphs and Line graphs.

The following table provides information on the number of National Lunch program participants in three states.

State	National School Lunch Program participants (all values are in millions)			
	2011	2012	2013	2014
California	3.28	3.33	3.29	3.24
Florida	1.65	1.65	1.64	1.66
New York	1.82	1.80	1.72	1.70

Source: US Department of Agriculture.

- Construct a bar graph showing the number of participants in the National Lunch program in all three states during 2011.
- Construct a line graph for the years 2011 to 2014 showing the number of participants in the National Lunch program in all three states.

2-2C Summarizing Quantitative variables – Frequency distributions and histograms

The following table contains a sample of individual student grades for my first year Principles of Macroeconomics class:

80	91	69	36	78	70	47
60	63	52	26	37	60	37
41	67	75	67	77	67	42
52	90	82	67	54	64	62

a) Calculate frequencies, relatively frequencies and percent frequencies.

Class	Frequency	Relative Frequency	Percent Frequency

b) Produce a corresponding histogram.

2-2D Summarizing Quantitative variables – Frequency distributions and Histograms

The following tables provide the share price of 28 companies listed on the Dow Jones Industrial Average stock index (marketwatch.com, May 17, 2016)

<u>Company</u>	<u>Share Price</u>	<u>Company</u>	<u>Share price</u>
Caterpillar	71	Pfizer	33
Dupont	64	IBM	148
JPMorgan	62	Wal-Mart Stores	65
Goldman Sachs	155	VISA	77
Exxon Mobile	89	Proctor and Gamble	80
Walt Disney	100	Cisco Systems	27
Nike	57	Intel Corp	30
Chevron	101	Boeing Co.	132
Apple	93	United Technologies	100
Johnson and Johnson	113	American Express Co.	63
General Electric	30	McDonalds Corp	128
UnitedHealth Group	129	Travelers Cos.	110
Verizon Communications	51	Microsoft Corp.	51
3M	168	Coca-Cola Co.	45
Merck & Co	54	Home Depot Inc.	132

a) Calculate frequencies, relative frequencies and percent frequencies.

Class	Frequency	Relative Frequency	Percent Frequency
0-29			
30-59			
60-89			
90-119			
120-149			
150+			

b) Show the corresponding histogram.

2-2E Summarizing Quantitative variables – Stem and Leaf

The following table contains a sample of individual student grades for my first year Principles of Macroeconomics class:

80	91	69	36	78	70	47
60	63	52	26	37	60	37
41	67	75	67	77	67	42
52	90	82	67	54	64	62

- a) Produce a stem and leaf display.

2-2F Summarizing Quantitative variables – Stem and Leaf

The following table contains a sample of fictional starting salaries for business students.

57,011	49,934	56,525	73,560	49,670
55,380	65,743	49,134	52,360	53,064
45,182	61,446	60,690	56,559	39,499
46,560	64,334	46,101	71,366	58,006

- a) Produce a stem and leaf display with a leaf value of 1000.

2-3A Summarizing Two variables – Crosstabulation

The following table fictional salaries for individuals in four professions at three different levels of experiences. Respond to the questions below.

		Experience		
		Low	Medium	High
Profession	Economist	44000	63000	66000
		39000	55000	72000
		48000	58000	87000
	Accountant	52000	54000	64000
		37000	63000	68000
		48000	68000	68000
	Engineer	72000	60000	83000
		53000	56000	72000
		41000	63000	96000
	Physicist	52000	52000	51000
		31000	61000	60000
		29000	55000	61000

- a) Produce a crosstabulation of the two variables: Profession and Salary.
- b) Calculate row and column percentages corresponding to part a).

		Salary				
		('000s of dollars)				
		20-39	40-59	60-79	80-100	Total
Profession	Economist					
	Accountant					
	Engineer					
	Physicist					
	Total					

2-3B Summarizing Two variables – Crosstabulation

The following table fictional salaries for individuals in four professions at three different levels of experiences. Respond to the questions below.

		Experience		
		Low	Medium	High
Profession	Economist	44000	63000	66000
		39000	55000	72000
		48000	58000	87000
	Accountant	52000	54000	64000
		37000	63000	68000
		48000	68000	68000
	Engineer	72000	60000	83000
		53000	56000	72000
		41000	63000	96000
	Physicist	52000	52000	51000
		31000	61000	60000
		29000	55000	61000

- a) Produce a crosstabulation of the two variables: Experience and Salary
- b) Calculate row and column percentages corresponding to part a).

		Salary (‘000s of dollars)				Totals
		20-39	40-59	60-79	80-100	
Experience	Low					
	Medium					
	High					
	Total					

2-4A Summarizing Two variables – Graphical Displays

The following table provides a number of pieces of information on a variety of different vehicles. Respond to the questions below.

Model	Engine size	Cylinders	Transmission	Number of gears	Fuel	Vehicle Class	City MPG	Hwy MPG	Model Year
ACURA RDX	3.5	6	Automatic	6	Gasoline	small SUV	19	28	2010
HYUNDAI Sonata	1.6	4	Manual	7	Gasoline	large car	28	38	2015
FORD Fiesta	1.6	4	Manual	5	Gasoline	small car	28	36	2006
DODGE Challenger	6.4	8	Automatic	8	Gasoline	midsize car	15	25	2010
BMW X5 xDrive35i	3	6	Automatic	8	Gasoline	standard SUV	18	24	2013
HONDA Accord	3.5	6	Automatic	6	Gasoline	midsize car	21	32	2014
LAND ROVER Range Rover	3	6	Automatic	8	Diesel	standard SUV	22	29	2008
BENTLEY Mulsanne	6.8	8	Automatic	8	Gasoline	midsize car	11	18	2012
MAZDA CX-5	2.5	4	Automatic	6	Gasoline	small SUV	24	30	2013
PORSCHE Cayman GTS	3.4	6	Manual	6	Gasoline	small car	19	26	2015

- Produce a scatterplot using City MPG on the y axis and Cylinders on the x axis.
- Add a trendline to your scatterplot and comment on any trend you observe.

Module 3: Numerical Summaries

3-1A Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following table provides a number of pieces of information on a variety of different vehicles. Respond to the questions below.

Model	Engine size	Cylinders	Transmission	Number of gears	Fuel	Vehicle Class	City MPG	Hwy MPG	Model Year
ACURA RDX	3.5	6	Automatic	6	Gasoline	small SUV	19	28	2010
HYUNDAI Sonata	1.6	4	Manual	7	Gasoline	large car	28	38	2015
FORD Fiesta	1.6	4	Manual	5	Gasoline	small car	28	36	2006
DODGE Challenger	6.4	8	Automatic	8	Gasoline	midsize car	15	25	2010
BMW X5 xDrive35i	3	6	Automatic	8	Gasoline	standard SUV	18	24	2013
HONDA Accord	3.5	6	Automatic	6	Gasoline	midsize car	21	32	2014
LAND ROVER Range Rover	3	6	Automatic	8	Diesel	standard SUV	22	29	2008
BENTLEY Mulsanne	6.8	8	Automatic	8	Gasoline	midsize car	11	18	2012
MAZDA CX-5	2.5	4	Automatic	6	Gasoline	small SUV	24	30	2013
PORSCHE Cayman GTS	3.4	6	Manual	6	Gasoline	small car	19	26	2015

- What is the mean, median and mode of City MPG?
- What is the mean, median and mode of Highway MPG?

3-1B Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following tables provide the share price of 30 companies listed on the Dow Jones Industrial Average stock index (marketwatch.com, May 17, 2016)

<u>Company</u>	<u>Share Price</u>	<u>Company</u>	<u>Share price</u>
Caterpillar	71	Pfizer	33
Dupont	64	IBM	148
JPMorgan	62	Wal-Mart Stores	65
Goldman Sachs	155	VISA	77
Exxon Mobile	89	Proctor and Gamble	80
Walt Disney	100	Cisco Systems	27
Nike	57	Intel Corp	30
Chevron	101	Boeing Co.	132
Apple	93	United Technologies	100
Johnson and Johnson	113	American Express Co.	63
General Electric	30	McDonalds Corp	128
UnitedHealth Group	129	Travelers Cos.	110
Verizon Communications	51	Microsoft Corp.	51
3M	168	Coca-Cola Co.	45
Merck & Co	54	Home Depot Inc.	132

- a) What is the mean, median and mode of DJIA companies share price?

3-1C Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following table provides average annual returns on two stocks over five years:

Year	Apple	Growth factor	Microsoft	Growth factor
2011	25.56%	1.26	-6.99%	0.93
2012	31.40%	1.31	2.89%	1.03
2013	5.42%	1.05	40.06%	1.40
2014	37.72%	1.38	24.16%	1.24
2015	-4.64%	0.95	19.44%	1.19

- a) Which stock had the highest average return over this five-year period?

3-1D Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following tables provide the share price of 30 companies listed on the Dow Jones Industrial Average stock index (marketwatch.com, May 17, 2016)

<u>Company</u>	<u>Share Price</u>	<u>Company</u>	<u>Share price</u>
Caterpillar	71	Pfizer	33
Dupont	64	IBM	148
JPMorgan	62	Wal-Mart Stores	65
Goldman Sachs	155	VISA	77
Exxon Mobile	89	Proctor and Gamble	80
Walt Disney	100	Cisco Systems	27
Nike	57	Intel Corp	30
Chevron	101	Boeing Co.	132
Apple	93	United Technologies	100
Johnson and Johnson	113	American Express Co.	63
General Electric	30	McDonalds Corp	128
UnitedHealth Group	129	Travelers Cos.	110
Verizon Communications	51	Microsoft Corp.	51
3M	168	Coca-Cola Co.	45
Merck & Co	54	Home Depot Inc.	132

- Calculate the 25th, 50th, and 75th quartiles.
- Calculate the 60th percentile.

3-1E Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following table contains a sample of individual student grades for my first year Principles of Macroeconomics class:

80	91	69	36	78	70	47
60	63	52	26	37	60	37
41	67	75	67	77	67	42
52	90	82	67	54	64	62

- What is the mean, median and mode of these grades?
- Compute the third quartile.
- Compute the 85th percentile.

3-1F Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following table is taken from my Principles of Microeconomics course outline. It explains the weights given to each of the course components when final grades are calculated.

Assignment	10%
Quizzes (3)	20%
Midterm Exams (2)	40%
Final Exam	30%

Before writing the final exam, you wish to calculate your interim grade. Your current grades on completed work are as follows:

Course component	Grade (%)
Assignment	62
Quiz #1	53
Quiz #2	65
Quiz #4	78
Midterm #1	67
Midterm #2	79

- a) What is your current interim grade?
- b) What grade must you receive on the final exam to receive a grade of 80 percent?

3-1G Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following data provides information on the distribution of administrator salaries at your university.

Salary Range (‘000s \$US)	Frequency
60-79	2
80-99	4
100-119	7
120-139	9
140-159	5
160-179	3

- a) Calculate the mean salary of these administrators.

3-2A Measures of Variability – Variance, Coefficient of Variation and Standard Deviation

The following table contains a reduced sample of individual student grades for my first year Principles of Macroeconomics class:

<u>Grade</u>
37
47
60
60
63
70
80
91

The sample mean is 63.5.

- Compute the range.
- Compute the interquartile range.
- Compute the variance.
- Compute the standard deviation
- Compute the coefficient of variation.

3-2B Measures of Variability – Variance, Coefficient of Variation and Standard Deviation

The following table contains CO2 emissions per capita for 10 EU member countries.

		CO2 per capita (metric tons of CO2)
	EU Countries	2010
1	Sweden	5.6
2	France	5.8
3	Spain	5.9
4	Italy	6.7
5	Greece	7.6
6	United Kingdom	8.0
7	Poland	8.3
8	Denmark	8.3
9	Germany	9.1
10	Netherlands	11.0

Source: Carbon Dioxide Information Analysis Center. <http://cdiac.ornl.gov/>

The sample mean is 7.6.

- Compute the range.
- Compute the interquartile range.
- Compute the variance.
- Compute the standard deviation
- Compute the coefficient of variation.

3-2C Measures of Location – Mean, Median, Mode, Percentiles and Quartiles.

The following data provides information on the distribution of administrator salaries at your university.

Salary Range (‘000s \$US)	Frequency
60-79	2
80-99	4
100-119	7
120-139	9
140-159	5
160-179	3

The mean salary is \$122,830.

- a) Calculate the variance and standard deviation.
- b) Calculate the coefficient of variation.

3-3A Measures of Shape and Relative Location

The following table contains CO2 emissions per capita for 10 EU member countries.

CO2 per capita <small>(metric tons of CO2)</small>	
EU Countries	2010
1 Sweden	5.6
2 France	5.8
3 Spain	5.9
4 Italy	6.7
5 Greece	7.6
6 United Kingdom	8.0
7 Poland	8.3
8 Denmark	8.3
9 Germany	9.1
10 Netherlands	11.0

Source: Carbon Dioxide Information Analysis Center. <http://cdiac.ornl.gov/>

The sample mean and variance are 7.6 and 2.7, respectively.

- Compute the z-score for Sweden, Greece, Denmark and Netherlands.
- Identify any countries that appear to be outliers using the first and third quartiles and the interquartile range.

3-3B Measures of Shape and Relative Location

The following data is a small sample of professors' salaries (in thousands of \$US) at your university.

<u>Salary</u> <u>('000s of \$US)</u>
46.0
71.6
75.1
83.1
84.0
87.1
92.2
93.5
95.3
<u>112.1</u>

- Determine the 1st and 3rd quartiles as well as the Interquartile range.
- Identify any outliers that may exist.

3-3C Measures of Shape and Relative Location

Upon graduation from college, you begin researching various labour markets in order to decide where you would like to move for work. You are able to find information on salaries of employees with 5 years of experience. In region A, the average salary in your field after 5 years is \$66,300 with a standard deviation of \$4,000. In region B, the average salary in your field after 5 years is \$69,000 with a standard deviation of \$7,000.

In both regions, the starting salary is \$55,000.

- a) Use Chebyshev's theorem and the Empirical Rule to analyze and compare these labour markets.

3-4A Measures of Relationship between two variables.

After writing a midterm exam, a small random sample of students was asked how many hours they studied in the two days preceding the exam. The following data contains their responses as well as the grade they received on the exam.

Hours studied (x_i)	Midterm grade (y_i)
2.6	75
3.8	78
4.7	90
5.2	81
1.6	70
$\bar{x} = 3.6$	$\bar{y} = 78.8$
$s_x = 1.5$	$s_y = 7.5$

- a) Calculate the covariance and the coefficient of correlation.

3-4B Measures of Relationship between two variables.

There exists a number of well-known websites where students have the opportunity to comment on their experience with their professors. In some cases, the grades are available for the students providing the evaluation. The following table contains a random sample of student evaluations of one instructor:

Student's grade (x_i)	Instructor's grade (y_i)
75	8
68	6
54	6
24	1
$\bar{x} = 55.3$	$\bar{y} = 5.3$
$s_x = 25.6$	$s_y = 3.0$

The student's grade is the grade the student received in the relevant professor's class, out of 100. The instructor's grade is the grade given by the corresponding student, out of ten.

- Calculate the covariance and the coefficient of correlation.

Module 4: Introduction to Probability

4-1A: Counting Rules and Probabilities

Suppose you are trying to plan your post-secondary education. Your parents are willing to pay your tuition if you remain within a 200 mile (320Km) radius of home. There exist two colleges within this radius. Because you aren't sure yet what you want to study, you're considering courses in either Arts or Science. Within each of these fields, each college offers three different majors. Consider this decision as an experiment.

- a) Develop a tree diagram for this experiment.
- b) List the experimental outcomes.

4-1B: Counting Rules and Probabilities

The game of Rock, Paper, Scissors has been resolving disagreements between siblings and friends for a millennium (actually traced back to China during the Han Dynasty 206BC-22-AD). If you need a refresher on the rules, see the article on Wikipedia.

- a) Develop a tree diagram for a two player game of Rock, Paper, Scissors.
- b) List the possible outcomes.
- c) What is the probability of winning this game?

4-1C: Counting Rules and Probabilities

On a hot summer day, you head down the street to a local ice cream shop. You can pay \$4 for two scoops, or \$5 for three scoops. There are 21 flavors from which to choose. How many different varieties of ice cream cones can you create if:

- a. You can have each flavor more than once?
- b. You can only have one scoop of each flavor?

(Note: The order in which the scoops are added counts ie. A scoop of chocolate on top of a scoop of vanilla is different than a scoop of vanilla on top of a scoop of chocolate.).

Consider scenario a) from above. If you asked the employee to guess what you want, what is the probability that he will make you exactly the ice cream cone that you want? What about in scenario b)?

4-1D: Counting Rules and Probabilities

Most passwords, either for your bank accounts, email account, or office computer, require a certain number of numbers, lower case letter and upper case letters.

- a. If your PIN number on your bank or credit card is required to be four digits long
 - i. How many possible PIN numbers are there? What is the probability of somebody guessing your PIN?
 - ii. How many would there be if you could only use each number once? What is the probability of somebody guessing your PIN?
- b. Most email passwords are required to be alpha-numeric. Let's assume yours must contain five characters.
 - iii. How many possible passwords are available using numbers and lower-case letters? What is the probability of somebody guessing your password? What if you could only use each number or letter once?
 - iv. How many possible passwords are available using numbers, lower-case and upper-case letters? What is the probability of somebody guessing your password? What if you could only use each number or letter once?

4-1E: Counting Rules and Probabilities

- 1) A vegetarian restaurant has a salad bar that offers 20 different ingredients. You can pay either \$5 to create a salad using five ingredients or \$8 for ten ingredients. Because the salad is going to be tossed and mixed together, it doesn't matter what order the ingredients are in. You are only allowed to use each ingredient once.
 - a. How many different kinds of salad can you make at each price?
 - b. If you asked your friend to guess what ingredients you would choose, what is the probability she would guess correctly?

- 2) After your healthy salad, you head down to the dessert bar. For a price of \$6 you can have any 3 desserts from a selection of 10. You can only choose each dessert once.
 - a. How many different combinations are there?
 - b. If you asked your friend to guess what ingredients you would choose, what is the probability she would guess correctly?

4-1F: Counting Rules and Probabilities

The chances of winning any lottery are slim. In the US, one of the largest lotteries is the PowerBall. Over time, the chances of winning this lottery have changed. In order to determine a winner prior to October 2015, officials would draw five numbered white balls from a drum of 59 and one numbered red ball from a drum of 35. This ball is known as the PowerBall. In order to win the jackpot, a player's ticket had to match the 5 numbers on the white balls, in any order, and the number from the red PowerBall. If only the 5 white ball numbers match, the player wins ONLY \$200,000.

On October 7th, 2015 the number of white balls in the drum increased to 69, while the number of red balls in the drum decreased to 26. The officials still draw five white and one red ball.

- a) Compute the number of different ways the first 5 numbers could be drawn before and after the change in October. What is the probability of winning the \$200,000 prize?
- b) Compute the total number of possible outcomes when the red ball is included, before and after the change in October. What is the probability of winning the jackpot?
- c) Imagine a scenario in which a player had to match not only the numbers on the five white balls, but must also have them in the correct order. How many possible outcomes would exist? What would be the probability of just matching the five white balls?

4-2A: Basics of Probability: Addition Law

A new pub has opened in town and you're considering it as a potential new spot for cheap wings and beer. Before going, you chat with a group of your friends who have already been. You obtain the following information:

- i) Three of them complained about the quality of the beer.
- ii) Four of them complained about the quality of the wings.
- iii) Two of them complained about the quality of the beer and wings.
- iv) Three of them had no complaints.

Answer the following questions:

- a) What is the probability you will enjoy the beer?
- b) What is the probability you will enjoy the wings?
- c) What is the probability you will dislike either the wings, or the beer, or both?
- d) What is the probability you will enjoy both the beer and the wings?
- e) Why have twelve of your friends already been here, but none have invited you?

4-2B: Basics of Probability: Addition Law

While registering for your courses, you find that the only option for your required course in economics is with Professor Dell. Because you've never had him as an instructor before, you decide to speak to some students who have. In the event of a lot of negative feedback, you may want to wait until next year and hopefully have more options. After talking with 13 students, you obtain the following information.

- i) Two students said Professor Dell doesn't seem to know the material.
- ii) Five students found Professor Dell to be unapproachable.
- iii) Three students said Professor Dell doesn't seem to know the material and unapproachable.

Given this feedback, respond to the following questions:

- a) What is the probability that Professor Dell is doesn't know the material?
- b) What is the probability that Professor Dell is unapproachable?
- c) What is the probability that Professor Dell either doesn't know the material, is unapproachable or both?
- d) What is the probability that Professor Dell is both knowledgeable and approachable?

4-2C: Basics of Probability: Addition Law

Your new business is in the process of upgrading their machinery and equipment. Prior to installation, it must go through a thorough quality assessment process by a third party inspector. In order to help plan for the necessary downtime, you request that the inspector provide you with a record of their previous inspections and how long the quality assessment process took. You are provided the following for the past ten inspections:

Number of days to completion	2	3	4	5	6	7
Number of inspections	1	3	2	2	1	1

Using this information, respond to the following questions:

- What is the probability that it will take 4 days or more to complete the inspection?
- What is the probability that it will take 5 days or fewer to complete the inspection?
- What is the probability that it will take 4 or 5 days?
- What is the probability that it will take 4 days or more, or 5 days or fewer?

4-3A: Basics of Probability: Conditional Probability and the Multiplication Law

A local alumni association performed a survey to determine labor market outcomes for their recent graduates in Business, Science and Arts. Specifically, they were asked whether or not they were able to obtain either full-time or part-time employment upon graduation. The following table contains the their responses:

	Business	Science	Arts	Total
Full-time	23	42	51	116
Part-Time	46	18	13	77
Total	69	60	64	193

- Develop a joint probability table for these data.
- Interpret the marginal probabilities.
- If a randomly selected person states they obtained part-time employment, what is the probability they studied Science?
- If a randomly selected person states they obtained full-time employment, what is the probability they studied Business?
- If a randomly selected person states that they studied Arts, what is the probability they obtained full-time employment?

4-3B: Basics of Probability: Conditional Probability and the Multiplication Law

As labor markets become increasingly competitive and individuals seek higher and higher levels of education, there are signs that these degrees may not be able to buy independence as easily as they once could. More and more students seem to be moving home after college. As a result, people are beginning to question whether or not a degree is necessary to gain independence. A sample of 517 people between the ages of 25 and 30 were asked whether or not they had a degree and whether or not they lived at home with their parents. The following table contains their responses:

	Live at home	Lives away	Total
Has a degree	109	187	296
Does not have a degree	89	135	224
Total	198	322	520

- f) Develop a joint probability table for these data.
- g) Interpret the marginal probabilities
- h) If a randomly selected person says they live at home, what is the probability that they have a degree?
- i) If a randomly selected person says they have a degree, what is the probability that they live at home?

4-3C: Basics of Probability: Conditional Probability and the Multiplication Law

The game of Rock, Paper, Scissors has been resolving disagreements between siblings and friends for a millennium (actually traced back to China during the Han Dynasty 206BC-22-AD). If you need a refresher on the rules, see the article on Wikipedia.

Answer the following questions; assuming only two people are playing the game:

- a) What is the probability that player one will win this game?
- b) What is the probability of either player one or player two winning this game?
- c) If player one wins the first game, what is the probability that she will win the second game?
- d) What is the probability that player one will win two games in a row?

4-4A: Bayes Theorem

Your first year economics class has so many students, that the instructor has hired 2 assistants, Jordan and Tony, to help with the marking. You ask a number of your friends who have taken this class before and find that roughly 95% of the assignments marked by Jordan receive passing grades, while only 85% of those marked by Tony receive a passing grade. It's safe to assume they each mark half of all the assignments.

- a) What is the probability of having Tony grade your assignment?
- b) What is the probability that if Tony graded your assignment, you will pass?
- c) What is the probability of having your assignment graded by Tony, and passing?
- d) Without knowing who marks your assignment, what is the probability of passing the assignment?
- e) What is the probability that if you pass the assignment, that it was Tony who a graded it?

4-4B: Bayes Theorem

You recently heard that a local restaurant has hired a highly talented and very well reviewed new head chef. Upon hearing the news, you go online to a restaurant review site and find that the current chef has a customer satisfaction rating of 80% while the new chef has a customer satisfaction rating of 98%. As part of the transition, to the new chef, the current chef is tasked with helping the new chef become familiar with his new kitchen. At the moment, you learn the current chef is preparing 75% of the meals, while the new chef is preparing 25% of the meals.

- a) What is the probability of having the new chef prepare your meal?
- b) What is the probability that if the new chef prepares your meal, you will be satisfied?
- c) What is the probability of having your meal prepared by the new chef and being satisfied?
- d) Without knowing who prepares your meal, what is the probability of being satisfied?
- e) After consuming your meal, you find it was delicious. What is the probability that it was prepared by the new chef?

Module 5: Discrete Probability Distributions

5-1A: Discrete and Continuous Random variables

For each of the following experiments, define a random variable and state whether the variable is discrete or continuous.

- a) Write your statistics exam, which is composed of 50 multiple choice questions.
- b) Observe the birds living near your favorite lake.
- c) Monitor production quality on an assembly line.
- d) Measure police response time to an emergency.
- e) Observe hourly traffic over a bridge.
- f) Monitor monthly crop yields.

5-2A: Discrete Probability distributions.

The following table shows a grade distribution for a typical first year economics course.

Letter Grade	Corresponding Grade Point	Number of students
A	4.00	9
B	3.00	11
C	2.00	21
D	1.00	1
F	0.00	8

Use this information to develop a probability distribution and to answer the following questions:

- Show that the probability distribution satisfies the conditions for a discrete probability distribution.
- What is the probability that a randomly selected student receives a grade of B or better?
- What is the probability that a randomly selected student receives a grade of D or worse?
- What is the probability that a randomly selected student receives a grade of B or C?

5-2B: Discrete Probability distributions.

Suppose your hometown is considering replacing the stop signs with traffic lights at one of its intersections. It first wants to determine how many cars pass through the intersection on a daily basis. The following table contains information on the number cars used the intersection over a 30 day period.

Number of cars	Number of days
1	5
2	6
3	9
4	6
5	4

Use this information to develop a probability distribution and to answer the following questions:

- a) Show that the probability distribution satisfies the conditions for a discrete probability distribution.
- b) What is the probability that on any given day three cars or less will use the intersection?
- c) What is the probability that on any given day, three cars or more will use the intersection?
- d) What is the probability that on any given day, only two or three cars will use the intersection?

5-3A: Expected Value and Variance

The following table shows a grade distribution for a typical first year economics course.

Letter Grade	Corresponding Grade Point	Number of students
A	4.00	9
B	3.00	11
C	2.00	21
D	1.00	1
F	0.00	8

Use this information to develop a probability distribution and to answer the following questions:

- Compute the expected value of a student's grade in this course?
- Compute the variance of grade points in the course?

5-3B: Expected Value and Variance

Suppose your hometown is considering replacing the stop signs with traffic lights at one of its intersections. It first wants to determine how many cars pass through the intersection on a daily basis. The following table contains information on the number cars used the intersection over a 30 day period.

Number of cars	Number of days
1	5
2	6
3	9
4	6
5	4

Use this information to develop a probability distribution and to answer the following questions:

- Compute the expected value for the number of cars using the intersection each day.
- Compute the variance of the number of cars using the intersection each day.

5-3C: Expected Value and Variance

Completing a doctoral degree can be a daunting task for anyone. Workload aside, the time commitment can be significant. However, not everybody takes the same amount of time. Suppose we randomly survey a group of 50 PhD holders and obtain the following information on the number of years they took to complete their doctorate.

Years	Number of students
2	4
3	7
4	15
5	14
6	10

Use this information to develop a probability distribution and to answer the following questions:

- Show that the probability distribution satisfies the conditions for a discrete probability distribution.
- What is the probability of completing a PhD in four years or less?
- What is the probability of completing a PhD in 4 years or more?
- Compute the expected value for the number years it takes to complete a PhD.
- Compute the variance of the number years it takes to complete a PhD.

5-4A: Bivariate Distributions and Covariance

Imagine you are asked to perform a study that considers the relationship, if any, between the price and quality of food at 100 restaurants. As part of this study, you determine three price levels (1 is inexpensive, 2 is moderate and 3 is expensive) and rate the quality of the meal on a scale of 1 to 3 (1 is poor, 2 is good and 3 is excellent). You obtain the following information:

Meal Price (y)	Meal Quality (x)			Total
	1	2	3	
1	7	3	5	15
2	9	24	7	40
3	4	23	18	45
Total	20	50	30	100

- Develop a bivariate probability distribution for quality and price at a randomly selected restaurant.
- Compute the expected value and variance for quality.
- Compute the expected value and variance for price.
- Compute the $var(x + y)$.
- Compute the covariance between price and quality.
- Compute the correlation coefficient between quality and price.

5-4B: Bivariate Distributions and Covariance

Your parents are interested in constructing an investment portfolio and have come to you for assistance. Two stocks are being considered. Let the expected value and variance of the first stock be given by: $E(x) = 4.1\%$ and $\text{var}(x) = 1.2$. Let the expected value and variance of the second stock be given by $E(y) = 9.4\%$ and $\text{var}(y) = 36$. The covariance between the two stocks is $\sigma_{xy} = -3$.

- a) Which of these two stocks is a riskier investment?
- b) What is the expected return and standard deviation, in dollars, for an investment of \$500 in the first stock?
- c) What is the expected return and standard deviation for an investment of 50 percent in each stock?
- d) What is the expected return and standard deviation for an investment of 75 percent in the first stock and 25 percent in the second?
- e) Compute the correlation coefficient between the returns of the two stocks.

5-5A: Binomial Probability Distribution

As the first student to arrive in class, you entertain yourself by applying some of what you've learned in statistics. You realize you can treat the gender of students entering the room as a binomial experiment. Let's assume the university's profile page states that its current student population consists of 40% female students and 60% male students.

- a) Compute the probability that your class of 20 students will contain exactly 5 females.
- b) Compute the probability that your class of 20 students will contain no females.
- c) Compute the probability that your class of 20 students will contain less than 5 females.
- d) What is the expected number of female students?
- e) What is the variance and standard deviation of the number of female students?

5-5B: Binomial Probability Distribution

The electronics retail industry typically has a high level of employee turnover, largely due to the fact that it is a common job for university students. Experience has shown that roughly 30 percent of new hires generally leave within 4 months of employment. Assume the manager hires ten students in August.

- a) What is the probability that at least 3 of the 10 new hires will quit within the first 4 months of employment?
- b) What is the probability that at least 5 of the 10 new hires will quit within the first 4 months of employment?
- c) What is the probability that none of the new hires quit within the first 4 months of employment?
- d) What is the expected value of the number of new hires who will quit in the first 4 months?
- e) What is the variance and standard deviation of the number of new hires who will quit with the first 4 months?

5-6A: Poisson Probability Distribution

We all despise line ups at grocery stores (or banks, or roller coasters, etc). The arrival of customers at the till follows a Poisson distribution. Assume that at your local grocery store, the average number of customer arrivals at the till is 6 every ten minutes.

- a) Compute the probability of no customers coming to the till during a ten minute period.
- b) Compute the probability of 2 or more customers coming to the till during a ten minute period.
- c) Compute the probability of no customers coming to the till during a five minute period.

5-6B: Poisson Probability Distribution

As internet providers race to gain a competitive advantage, more and more are switching their networks to fiber optic cable. The conversation requires the installation of thousands of miles of optical cables. Most of the time, the installation goes without any problems, however, the fiber optic cables are never perfectly flawless. In order to ensure optimum performance for their customers, the installers must repair or replace any sections of the cable that are found to be defective. Assume the cables show an average of 2 defects per 10 miles of cable.

- a) Compute the probability of finding zero defects in 20 miles of cable.
- b) Compute the probability of finding one defect in 20 miles of cable.
- c) Compute the probability of finding two or more defects in 20 miles of cable.
- d) Compute the probability of finding exactly 2 defects in 20 miles of cable.

5-7A: Hypergeometric Probability Distribution

A shipping company receives a shipment of 15 items. A random sample of 6 is selected, and the shipment is rejected if any of these items fails to meet the minimum quality specifications.

- a) What is the probability of accepting a shipment with 5 defective items?
- b) What is the probability of accepting a shipment with 1 defective item?
- c) What is the probability of rejecting a shipment with 1 defective item?

5-7B: Hypergeometric Probability Distribution

Black Jack is a popular game in casinos and the homes of those who enjoy gambling. A player is dealt two cards. Face cards (jacks, queens and kings) and tens are worth 10 points. Aces can be scored as either a one or eleven. All other cards are scored at face value. A 52 card deck contains 16 cards worth 10 and four aces. The best hand possible, called BlackJack, is a hand with 21 points. This would consist of one ace and one 10 point card.

- a) What is the probability that both cards are aces or 10-point cards?
- b) What is the probability that both of the cards are aces?
- c) What is the probability that both of the cards are 10-point cards?
- d) What is the probability that a player is dealt a BlackJack?

Module 6: Continuous Probability Distributions

6-1A: Uniform Probability Distribution

Assume a random variable x is known to be uniformly distributed between 2 and 5.

- a) Draw the graph of the probability density function.
- b) Compute $P(x = 2.8)$
- c) Compute $P(2.5 \leq x \leq 3.5)$
- d) Compute $P(3 < x < 4.2)$
- e) Compute $E(x)$ and $\text{var}(x)$

6-1B: Uniform Probability Distribution

Anybody who has taken university courses knows how instructors can often veer off topic during their lectures. Let's assume that during a 50 minute class, the instructor spend an average of 10 minutes talking about something completely unrelated to the course. Let's assume further that that amount of time spent lecturing is uniformly distributed between 35 and 45 minutes.

- a) Draw the graph of the probability density function.
- b) What is the probability that a class will have 42 minutes or more of relevant lecturing?
- c) What is the probability that a class will have between 38 and 42 minutes of relevant lecturing?
- d) What is the probability that a class will have less than 40 minutes of relevant lecturing?
- e) What is the probability that the professor spends more than 12 minutes of topic?

6-2A: Normal Probability Distribution

Midterm exams are always a challenge for my students. In addition to the difficult content, time constraints are often binding for many. As an instructor, it's important that my exams be long enough to cover all of the relevant material, but short enough for students to be able to complete in the given time. For a 75 minute class, I write an exam that I hope will be completed in an average of 65 minutes. Assume the distribution of completion times is a normally distributed variable. For the most recent exam, the average completion time was 64 minutes with a standard deviation of 7.28 minutes. Answer the following questions.

- a) What is the probability of completing the exam before the end of class (75 minutes or less)?
- b) What is the probability that a student will fail to complete the exam in before the end of class?
- c) What is the probability that a student will complete the exam in more than 60 minutes but less than 75 minutes?
- d) If the class contains 30 students, how many will fail to complete the exam before the end of class?

6-2B: Normal Probability Distribution

Whether you're a student or a professional, attending meetings will always be a fact of life. During my student and professional life, it's safe to say I have never attended a meeting that has started on time. Assume the time relative to a meeting's start time is a normally distributed variable. The variable takes the value of zero when the meeting starts on time, is positive when it starts late and is negative when it starts early. In my current employment, meetings seem to start an average of 10 minutes late, with a standard deviation of 3.9 minutes.

- a) What is the probability that a meeting starts exactly on time?
- b) What is the probability that a meeting starts early?
- c) What is the probability that a meeting starts late?
- d) What is the probability that a meeting starts at least 20 minutes late?

6-2C: Normal Probability Distribution

Given that z is a standard normal random variable, compute the following probabilities:

- a) $P(z \leq 1)$
- b) $P(z \leq 0.5)$
- c) $P(z \leq -1)$
- d) $P(-1 < z \leq 1.5)$

6-3A: Normal Approximation of Binomial Probabilities

A binomial probability distribution has $p = 0.35$ and $n = 80$.

- a) What are the mean and standard deviation?
- b) Given the parameters of this distribution, can it be approximated by the normal probability distribution? Explain.
- c) What is the probability of exactly 20 successes?
- d) What is the probability of 14 to 18 successes?
- e) What is the probability of 16 or fewer successes?

6-3B: Normal Approximation of Binomial Probabilities

According to the 2014 General Social Survey (GSS), a project by the independent research organization NORC at the University of Chicago, although the number of guns per capita in the United States is on the rise (112.6 guns per 100 residents), the number of households with at least one gun, has declined to roughly 1 in 3. The implication being, there are a growing number of guns concentrated into fewer homes. Consider a group of 300 homes.

- a) What is the expected number of households that own at least one gun?
- b) What is the probability that 90 or fewer own a gun?
- c) What is the probability that between 85 and 115 own a gun?
- d) What is the probability that 115 or more own a gun?

6-4A: Exponential Probability Distribution

Consider an exponential probability function with a mean of $\mu = 9$.

- a) What is the formula for $P(x \leq x_0)$?
- b) Compute $P(x \leq 7)$.
- c) Compute $P(x \leq 4)$.
- d) Compute $P(x \geq 11)$.
- e) Compute $P(7 \leq x \leq 11)$.

6-4B: Exponential Probability Distribution

Assume the time between arrivals of customers at a grocery store checkout follows an exponential probability distribution with a mean time of 6 minutes.

- a) What is the probability that the arrival time between customers will be 4 minutes or less?
- b) What is the probability that the arrival time between customers will be more than 7 minutes?
- c) What is the probability that the arrival time between customers will be between 4 and 7 minutes?

6-4C: Exponential Probability Distribution

With the growth in the use of emails and text messaging, it has become increasingly difficult to focus on one task for any length of time without constant interruptions. Assume that as a young and popular undergraduate student, your phone is chirping an average of 20 times per hour because of text messages from your friends. Assume the number of texts follows a Poisson probability distribution.

- a) Show the probability distribution for the time between texts.
- b) What is the probability you will receive no text messages during a 15 minute period?
- c) What is the probability that the next text message will be received within 5 minutes?

Module 7: Sampling and Sampling Distributions

7-1A: Selecting a sample

Consider a finite population with four elements labelled A, B, C and D. There exist six possible simple random samples of size $n=2$ that can be selected.

- a) List the six samples.
- b) Using simple random sampling, what is the probability that each of the $n=2$ samples is selected?

7-1B: Selecting a sample

Consider a small town that consists of five households. The following data list the number of children per household.

Household	Children
A	2
B	1
C	4
D	3
E	4

- List all possible samples of $n = 2$, without replacement.
- List all possible samples of $n = 4$, without replacement.
- Using simple random sampling, what is the probability that each of the $n = 2$ samples is selected?
- Using simple random sampling, what is the probability that each of the $n = 4$ samples is selected?

7-2A: Point Estimation

Consider the following data taken from a simple random sample: 4 8 3 9 13 14

- a) What is the point estimate of the population mean?
- b) What is the point estimate of the population standard deviation?

7-2B: Point Estimation

Consider a small town that consists of five households. The following data list the number of children per household. The population average number of children per household is $\mu = 2.8$ with a standard deviation of $\sigma = 1.3$.

Household	Children
A	2
B	1
C	4
D	3
E	4

- g) Using the simple random sampling method, a sample consisting of households C and D is taken. Calculate the point estimate of the mean and standard deviation.
- h) Using the simple random sampling method, a sample consisting of households A, B, D and E is taken. Calculate the point estimate of the mean and standard deviation.

7-2C: Point Estimation

A recent survey by the Pew Research Center asked registered voters ($n=1,655$) in the US about the importance of various issues in deciding who they would vote for in the 2016 election. (*2016 Campaign: Strong Interest, Widespread Dissatisfaction*, July 7, 2016).

- a) What is the sampled population for this study?
- b) The Economy was cited as “very important” by 84% of respondents. Estimate the number of respondents who believe this issue is very important.
- c) Health care was cited as “very important” by 1,225 respondents. Estimate the proportion of registered voters in the US who believe this issue is very important.
- d) Foreign Policy was cited as “very important” by 1,241 respondents. Estimate the proportion of registered voters in the US who believe this issue is very important.
- e) What is the target population for the inferences in parts c) and d)? Is it the same as the sampled population identified in part a)? Would it be accurate to extend these inferences to the entire U.S. population? Why or why not?

7-3A: Sampling Distributions

A given population has a mean of 300 and a standard deviation of 40. A sample of size $n = 125$ is taken.

- a) What is the expected value of \bar{x} ?
- b) What is the standard deviation of \bar{x} ?
- c) Describe the sampling distribution of \bar{x} .
- d) What does this sampling distribution show?

7-3B: Sampling Distributions

My home university has a population of $N = 25,748$ students with an average age of $\mu = 25.1$. If we select a simple random sample of 200 students,

- a) Is the finite population correction factor necessary in calculation the standard error?
- b) Using a population standard deviation of $\sigma = 5.36$, calculate the standard error both with and without the correction factor. Discuss the reasoning behind ignoring the correction factor when $n/N \leq 0.05$.
- c) What is the probability that the sample mean age will be within 0.5 years of the population mean age?

7-3C: Sampling Distributions

According to an article by Time Magazine, starting salaries by undergraduate majors varies quite significantly. The following table provides the top 5 projected salaries by discipline for 2016:

Area of study	2016 average salary
Engineering	\$64,891
Computer Science	\$61,321
Math and Sciences	\$55,087
Business	\$52,236
Agriculture/Natural Resources	\$48,729

Source: <http://time.com/money/3829776/heres-what-the-average-grad-makes-right-out-of-college/> (retrieved July 2016)

Assume a population standard deviation of $\sigma = \$6,410$

- i) What is the probability a sample of 100 Engineering majors, will provide a sample mean starting salary within \$750 of the population mean?
- j) What is the probability a sample of 100 Computer Science majors, will provide a sample mean starting salary within \$750 of the population mean?
- k) What is the probability a sample of 150 Math and science majors, will provide a sample mean starting salary within \$750 of the population mean? Discuss the differences you observe between this probability and the ones calculated in parts a) and b).

7-4A: Sampling Distributions (\bar{p})

My home university has a student population of 25,748 students. Of those, 2,635 are international students. A sample of 100 students is selected.

- a) What is the expected value of \bar{p} ?
- b) What is the standard error of \bar{p} ?
- c) What is the probability that the sample proportion will be within 0.05 of the population proportion?
- d) If the sample size is 300, what is the probability that the sample proportion will be within 0.05 of the population proportion?

7-4B: Sampling Distributions (\bar{p})

A population proportion is 0.45. A sample of $n = 150$ is taken.

- a) What is the probability that the sample proportion is within 0.02 of the population proportion?
- b) What is the probability that the sample proportion is within 0.04 of the population proportion?

Module 8: Estimating Intervals

8-1A: Interval estimates of a population mean

A simple random sample of 35 items has a sample mean of 120. The standard deviation is known to be $\sigma = 24$.

- a) Calculate the standard error of the mean.
- b) Calculate the margin of error for a 95% confidence interval.
- c) Calculate the limits of the 95% confidence interval.

8-1B: Interval estimates of a population mean

According to an article by Time Magazine, starting salaries by undergraduate majors varies quite significantly. The following table provides the top 5 projected salaries by discipline for 2016:

Area of study	2016 average salary
Engineering	\$64,891
Computer Science	\$61,321
Math and Sciences	\$55,087
Business	\$52,236
Agriculture/Natural Resources	\$48,729

Source: <http://time.com/money/3829776/heres-what-the-average-grad-makes-right-out-of-college/> (retrieved July 2016)

Assume a population standard deviation of $\sigma = \$6,410$

- Using a sample of size $n = 100$, calculate a 95% confidence interval for the average salary of an Engineering major.
- Using a sample of size $n = 100$, calculate a 99% confidence interval for the average salary of an Engineer.
- Using a sample of size $n = 150$, calculate a 95% confidence interval for the average salary of an Engineer.

8-1C: Interval estimates of a population mean

The following sample data are taken from a normally distributed population:

2 6 11 5 13 7

- a) Calculate the point estimate of the population mean.
- b) Calculate the point estimate of the population standard deviation.
- c) Determine the t value corresponding to an area of 0.025 with the appropriate degrees of freedom.
- d) Calculate the limits of a 95% confidence interval for the population mean.

8-1D: Interval estimates of a population mean

A simple random sample of households was taken to study their electricity usage. Their monthly electricity consumption was recorded as follows (measure in Kilowatt-hours):

A	B	C	D	E
320.4	295.6	333.2	287.9	302.8

- l) Calculate the point estimate of the population mean.
- m) Calculate the point estimate of the population standard deviation.
- n) Determine the t value needed for a 90% confidence interval estimate of the population mean.
- o) Calculate the limits of a 90% confidence interval for the population mean and interpret the results.

8-2A: Interval estimates of a population proportion

Upon realizing you have 1296 Facebook friends, you decide to do a quick check to determine how many of them are actually your friends in real life. Rather than check all 1296, you take a random sample of 200. Of those, you find that only 23 are people with whom you would actually want to spend time. The rest must just be friends of friends or people who have added you for some reason.

- a) What is the point estimate of the population proportion of your Facebook friends who are actually real world friends?
- b) Calculate the standard error of \bar{p} .
- c) Calculate a 95% confidence interval for the proportion of your Facebook friends who are actually real world friends.

8-2B: Interval estimates of a population proportion

Radar traps are frequently posted in locations where they are most likely to catch people driving over the posted speed limit. In one location, police officers said that 124 out of 200 vehicles were charged with speeding.

- a) What is the point estimate of the population proportion of speeders at this location?
- b) Calculate the standard error of \bar{p} .
- c) Calculate a 95% confidence interval for the population proportion of speeders at this location.
- d) Calculate a 90% confidence interval for the population proportion of speeders at this location.