

LEARNING EXCEL : DATA ANALYSIS

An Internship Project Report Submitted in Partial
fulfillment of the requirements for the award of
Degree of Bachelor of Science in Mathematics

Submitted by

A.Pavithradevi

Register No: 18BM7462

Under the guidance of

Mrs. D. Kavitha M.Sc., M.Phil.,

Assistant Professor

Department of Mathematics (Self Finance)



Sri G.V.G Visalakshi College for Women
(Autonomous)

(Affiliated to Bharathiar University, Coimbatore)

Accredited at 'A⁺' Grade by NAAC (CGPA 3.27)

An ISO 9001:2015 Certified Institution

Udumalpet-642 128

March 2021



LinkedIn LEARNING

Certificate of Completion
Congratulations, Pavithradevi A

Learning Excel: Data Analysis

Course completed on Mar 09, 2021 at 02:14PM UTC • 2 hours 9 min

By continuing to learn, you have expanded your perspective, sharpened your skills, and made yourself even more in demand.

A handwritten signature in black ink that reads "David Roberts".

Head of Content Strategy, Learning

LinkedIn Learning
1000 W Maude Ave
Sunnyvale, CA 94085

Certificate Id: AZQG5sD2oGm_Gla9sw3nliPwEv9g

CERTIFICATE

This is to certify that the internship project work entitled **“LEARNING EXCEL: DATA ANALYSIS”** is a bonafied record done by **A.PAVITHRADEVI (18BM7462)** submitted in partial fulfillment of the requirement for the award of the degree of Bachelor of Science in Mathematics at Sri G.V.G Visalakshi college for women (Autonomous), Udumalpet during the academic year 2018-2021.

Signature of the HOD

Signature of the staff-in-charge

Introduction

The internship training program was organized by “Skill India- eskillIndia E-learning aggregator from NSDC (National Skill Development Corporation). This is one of the E-Learning Content. This E skill India include many Knowledge Partners. One of the Knowledge Partner “LinkedIn” gave this Course. Author **Curt Frye** teaches about the foundation concepts including basic calculations such as mean, median, mode, variance and standard deviation and also provides how to calculate the covariance and correlation and utilize data distributions.

Digital Tools



Dr. Curt Frye gave training under the title LEARNING EXCEL : DATA ANALYSIS

1. Introduction about the power data analysis of using excel
2. To calculate mean, median and mode
3. To measure maximum and minimum
4. To find the Error and Squared error from the mean value
5. Analyze data using variance and standard deviation
6. Use of Normal distribution
7. Use of Poisson distribution
8. Use of Exponential distribution
9. To Measure Covariance
10. To Measure Correlation

Tools used for Learning Excel: Data Analysis



Introduction about the power of data analysis of using excel

- ❖ Returns the correct results about 85% of the time.
- ❖ Time saving
- ❖ Gather and organize the data effectively.
- ❖ Create formulas in excel.
- ❖ Create charts.

To calculate Mean, Median and Mode

❖ Mean formula

1. =average(a2:a12)
2. Where, a2:a12 is the column A entered data
3. Finally, click on enter to get the Mean value from the column A from a2 to a12.

❖ Median formula

1. =median(a2:a12)
2. Where, a2:a12 is the column A entered data.
3. Finally, click on enter to get the Median value from the column A from a2 to a12.

❖ Mode formula

1. =mode.sngl(a2:a12)

Where, a2:a12 is the column A entered datas

Click enter to get mode.sngl value.

2. =mode.mult

Where, a2:a12 is the column A entered datas

Click ctrl+shift+enter to get mode.mult value.

Calculation of mean, median and mode:

The screenshot displays a LinkedIn Learning course page. On the left is a 'Contents' sidebar with a list of lessons. The main area shows a video player with an Excel spreadsheet titled 'MeanAndMedian - Excel'. The spreadsheet contains the following data:

Days in Transit			
23	Mean		29.4545
38			
29	Median		28
28			
30	Mode		=mode
28			
29			
27			
27			
28			
37			

The video player interface includes an 'Overview' tab, 'Q&A', 'Notebook', and 'Transcript' options. Below the video, the instructor is identified as Curt Frye, President of Technology and Society. A 'Give feedback' button is also visible.

To measure maximum and minimum

❖ Minimum formula

1. `=min(a2:a41)`
2. Where, a2:a41 is the column A entered datas.
3. Finally, click on `ctrl+shift+↓` to get all the values in column A.
4. Then, click on enter to get the minimum value from the column A from a2 to a41.

❖ Maximum formula

1. =average(a2:a41)
2. Where, a2:a41 is the column A entered datas
3. Finally, click on ctrl+shift+↓ to get all the values in column A.
4. Then, click on enter to get the minimum value from the column A from a2 to a41.

❖ Inclusive and exclusive values

Inclusive formula End with .inc

Exclusive formula end with .exc

The screenshot displays a LinkedIn Learning course interface. On the left, a 'Contents' sidebar lists several lessons, with 'Measure maximums, minimums, and other data characteristics' selected. The main content area shows a video player with an Excel spreadsheet. The spreadsheet has the following data:

Order Value	Minimum	Maximum	Inclusive	Exclusive
6,231	=min(A2)			
6,885				
9,637				
8,060				
7,265				
5,629				
1,861				
6,062				
6,286				

A tooltip indicates the action: 'select the next nonblank cell ctrl + shift + ↓'. Below the video player, the instructor's name 'Curt Frye' is visible, along with a 'Give feedback' button. The bottom of the screen shows the Windows taskbar with the search bar and system tray.

To find the Error and Squared errors from the Mean value

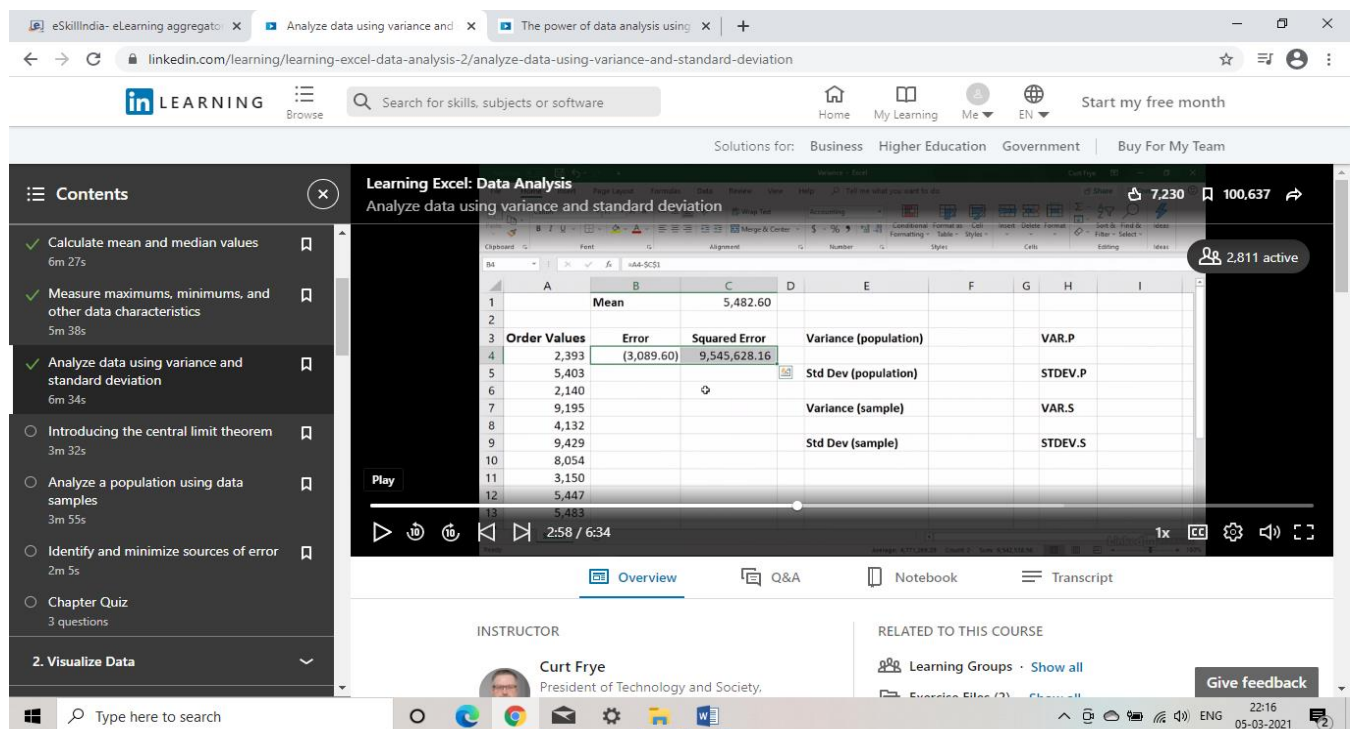
➤ Error finding:

- Error for b4
- Formula: $=a4-c1$ and click enter.
- Where, c1 is the entered mean value
(here the mean value taken as 5,482.60)
- Finally we get the “b4” error value for the order value 2,393 as 3,089.60

➤ Squared errors:

- Squared Error for “b4” is
- Formula : $=b4^2$ and click enter.
- Finally, the squared error is given as 9,545,628.16

Calculation of Error and Squared error:



The screenshot shows a LinkedIn Learning video player displaying an Excel spreadsheet. The spreadsheet is titled "Learning Excel: Data Analysis" and "Analyze data using variance and standard deviation". The data is as follows:

Order Values	Mean	Error	Squared Error	Variance (population)	Std Dev (population)	Variance (sample)	Std Dev (sample)
2,393	5,482.60	(3,089.60)	9,545,628.16				
5,403							
2,140							
9,195							
4,132							
9,429							
8,054							
3,150							
5,447							
5,483							

The video player interface includes a sidebar with a table of contents, a play button, and a progress bar. The instructor is identified as Curt Frye, President of Technology and Society. The video has 2,811 active viewers and is part of a course with 100,637 views.

Analyze data using variance and standard deviation

Variance:

- Variance population Formula: =average(c4:c13)
- Click enter to get variance in f3 from c4 to c13 (we calculate for squared errors values) we get 6,393,943.44
- And to use formula: =var.p(a4:a14) taken from order values we get as 6,393,943.44
- Thus, we get the result as the variation is as same as taken from squared error and order values

Standard deviation:

- Standard deviation Formula: =sqrt(f3)
- Click enter to get standard deviation in f5 from squaring f3 (we calculate from squared values) we get as 2,529.02
- And to use formula: =stdev.p(a4:a14) taken from order values we get as 2,529.02
- Thus, we get the result as the standard deviation is as same as taken from squared error and order values

The screenshot shows a LinkedIn Learning course interface. On the left is a 'Contents' sidebar with several lessons, including 'Analyze data using variance and standard deviation' which is currently selected. The main area displays a video player showing an Excel spreadsheet. The spreadsheet contains the following data:

	A	B	C	D	E	F	G	H	I
1		Mean		5,482.60					
2									
3	Order Values	Error	Squared Error		Variance (population)	6,395,943.44	VAR.P		6,395,943.44
4	2,393	(3,089.60)	9,545,628.16						
5	5,403	(79.60)	6,336.16		Std Dev (population)	2,529.02	STDEV.P		2,529.02
6	2,140	(3,342.60)	11,172,974.76						
7	9,195	3,712.40	13,781,913.76		Variance (sample)	7,106,603.82	VAR.S		7,106,603.82
8	4,132	(1,350.60)	1,824,120.36						
9	9,429	3,946.40	15,574,072.96		Std Dev (sample)	2,665.82	STDEV.S		2,665.82
10	8,054	2,571.40	6,612,097.96						
11	3,150	(2,332.60)	5,441,022.76						
12	5,447	(35.60)	1,267.36						
13	5,483	0.40	0.16						

Below the video player, there is an 'INSTRUCTOR' section for Curt Frye and a 'RELATED TO THIS COURSE' section. The bottom of the screen shows a Windows taskbar with the time 20:43 and date 05-03-2021.

- To calculate **variance sample** use formula:

$$=sum(c4:c13)/(count(c4:c13)-1$$
- This value changes from variance (variance population) because of -1
- We get same ans using $=var.s(a4:a13)$ click enter.
- To calculate **standard deviation** sample use formula: $=sqrt(f3)$ and same value using formula as $=stdev.s(a4:a13)$
- Click enter.

Data Distributions

- **Use of Normal distribution:**

NORM.DIST function

Syntax:

NORM.DIST(x, mean, standard_dev, cumulative)

The NORM.DIST function syntax has the following arguments:

- **X** - the value for which you want the distribution.
- **Mean** - the arithmetic mean of the distribution.
- **Standard_dev** - the standard deviation of the distribution.
- **Cumulative** Required. A logical value that determines the form of the function. If cumulative is TRUE, NORM.DIST returns the cumulative distribution function; if FALSE, it returns the probability density function.

Remarks:

- If mean or standard_dev is non-numeric, NORM.DIST returns the #VALUE! error value.
- If standard_dev ≤ 0 , NORM.DIST returns the #NUM! error value.
- If mean = 0, standard_dev = 1, and cumulative = TRUE, NORM.DIST returns the standard normal distribution, NORM.S.DIST.
- The equation for the normal density function (cumulative = FALSE) is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- When cumulative = TRUE, the formula is the integral from negative infinity to x of the given formula.

Normal distribution:

The screenshot displays a LinkedIn Learning video player. The video content shows an Excel spreadsheet with a normal distribution curve. The spreadsheet data is as follows:

Mean	Standard Deviation	Probability of exactly 84	Probability of 84 or less	71% of values below	85% of values below
100	15	0.01505752	0.14306119	108.301	115.547
36	2.9633E-06				
38	5.1877E-06				
40	8.922E-06				
42	1.5074E-05				
44	2.5019E-05				
46	4.0793E-05				
48	6.5342E-05				
50	0.00010282				
52	0.00015894				
54	0.00024136				
56	0.00036007				
58	0.0005277				
60	0.00075973				
62	0.00107452				

The normal distribution curve is titled "Normal Curve $\mu = 100, \sigma = 15$ ". The x-axis ranges from 0 to 180, and the y-axis ranges from 0 to 0.03. The curve is centered at 100. Below the spreadsheet, there is a search bar for the transcript and a "Give feedback" button.

- The normal distribution is defined by two values. The first is the mean(average) and the second is the standard deviation. The standard deviation is the amount of spread in your data.

Procedure:

- To create calculations for a variety of values. And those will appear in the chart that you see on the worksheet. Starting in cell B4, we calculate the probability of the value 36. So, in cell B4, =NORM.DIST. This takes a look at the probability of a specific value occurring within the normal distribution.
- This function requires four arguments. The first is the value in cell A4, as 36 and then a comma. The next is the mean, and in cell B1 then a comma. The standard deviation is in B2. And finally indicate whether the cumulative distribution function, that is, any value up

to, and including a specific value X . Or whether the point probability using the probability mass function. That is, the probability that a specific value will occur.

- To use the probability mass function, so highlight false, press tab.
- Click enter, the first value will display. And it doesn't even show on the chart. Now, copy this formula down for all of the cells in Column A that contain values. So click cell B4, double click the fill handle. Double click, and there you see the familiar shape of the normal curve.
- The average, or the peak occurs right at 100, and you can see that the values spread out and that is a function of the standard deviation of 15.
- To calculate the specific value, of say, 84 occurring, then click in cell E1 enter the formula for normal distribution in excel then click enter. We get the value about 0.15%
- Follow the same process to go for the probably of 84 or less, we have TRUE for the cumulative probability. Type right parenthesis and press Tab
- We get a value of 84 or less will happen about 14.3% of the time.
- To find the cutoff, where 71% of values will occur below a specific point. In this case, let's do 71% , we see that 71% of values would be below 108.3.
- The same thing for 85, we see that 85% of the values will occur below 115.5. Note that 14% of values will occur between 115.5 and 108.3. And if you look at the chart again, you will see that the normal curve is higher, has higher probability toward the middle, around 100, and that even at 115. So if we wanted to see the value, where we had 95 or even 98% of values, we'd be much further over to the right.

Use of Poisson distribution:

POISSON.DIST function

Returns the Poisson distribution. A common application of the Poisson distribution is predicting the number of events over a specific time, such as the number of cars arriving at a toll plaza in 1 minute.

Syntax:

POISSON.DIST(x,mean,cumulative)

The POISSON.DIST function syntax has the following arguments:

- **X** - the number of events.
- **Mean** - the expected numeric value.
- **Cumulative** - a logical value that determines the form of the probability distribution returned. If cumulative is TRUE, POISSON.DIST returns the cumulative Poisson probability that the number of random events occurring will be between zero and x inclusive; if FALSE, it returns the Poisson probability mass function that the number of events occurring will be exactly x.

Remarks:

- If x is not an integer, it is truncated.
- If x or mean is nonnumeric, POISSON.DIST returns the #VALUE! error value.
- If $x < 0$, POISSON.DIST returns the #NUM! error value.
- If $\text{mean} < 0$, POISSON.DIST returns the #NUM! error value.
- POISSON.DIST is calculated as follows.

If cumulative = FALSE:

$$\text{Poisson} = \frac{e^{-\lambda} \lambda^x}{x!}$$

If cumulative = TRUE:

$$\text{Compoission} = \sum_{k=0}^x \frac{e^{-\lambda} \lambda^k}{k!}$$

Example:

Copy the example data in the following table, and paste it in cell A1 of a new Excel worksheet. For formulas to show results, select them, press F2, and then press Enter. If you need to, you can adjust the column widths to see all the data.

Data	Description	
2	Number of events	
5	Expected mean	
Formula	Description	Result
=POISSON.DIST(A2,A3,TRUE)	Cumulative Poisson probability with the arguments specified in A2 and A3.	0.124652
=POISSON.DIST(A2,A3,FALSE)	Poisson probability mass function with the arguments specified in A2 and A3.	0.084224

Use of exponential function:

EXPON.DIST function

Returns the exponential distribution. Use EXPON.DIST to model the time between events, such as how long an automated bank teller takes to deliver cash. For example, you can use EXPON.DIST to determine the probability that the process takes at most 1 minute.

Syntax:

EXPON.DIST(x,lambda,cumulative)

The EXPON.DIST function syntax has the following arguments:

- **X** - the value of the function.
- **Lambda** - the parameter value.
- **Cumulative** - a logical value that indicates which form of the exponential function to provide. If cumulative is TRUE, EXPON.DIST returns the cumulative distribution function; if FALSE, it returns the probability density function.

Remarks:

- If x or lambda is nonnumeric, EXPON.DIST returns the #VALUE! error value.
- If $x < 0$, EXPON.DIST returns the #NUM! error value.
- If $\lambda \leq 0$, EXPON.DIST returns the #NUM! error value.
- The equation for the probability density function is:

$$f(x; \lambda) = \lambda e^{-\lambda x}$$

- The equation for the cumulative distribution function is:

$$f(x; \lambda) = 1 - e^{-\lambda x}$$

Example:

Copy the example data in the following table, and paste it in cell A1 of a new Excel worksheet. For formulas to show results, select them, press F2, and then press Enter. If you need to, you can adjust the column widths to see all the data.

Data	Description	
0.2	Value of the function	
10	Parameter value	
Formula	Description	Result
=EXPON.DIST(A2,A3,TRUE)	Cumulative exponential distribution function	0.86466472
=EXPON.DIST(0.2,10,FALSE)	Probability exponential distribution function	1.35335283

To Measure Covariance:

I. Covariance formula = $\frac{\sum(x-\bar{x})(y-\bar{y})}{n}$

Procedure:

- For each data point, find its deviation from the Mean.
- Multiply the deviation for each pair of data points.
- Find the sum of all these values and divide by the number of data pairs.

Interpreting covariance value:

- 0** - The data sets don't vary together.
- Positive** - The data sets tend to move in the same directions.
- Negative** - The data sets tend to move in the opposite directions.

Calculate Covariance between Two Columns of data:

- Labeled column one and column two. Covariance is the average of all of the differences or variances for pairs of data points.
- Type in `=(a2-average(a2:a11))*(b2-average(b2:b11))`
Press enter. We get the average of column A and column B
 - To find the average of all those covariance. And that is exactly the same as dividing by the number of data pairs.
 - Type `=AVERAGE(c2:c11)` and click enter we get the covariance as -0.48, which is very small.
 - To find between two columns type `=covariance.p(a2:a11,b2:b11)`
 - Click enter we get the value as - 0.48.
 - And also type `=covariance.s(a2:a11,b2:b11)` then click enter we get, -0.533333 a difference covariant function value. The difference is due to subtracting one from the number of data pairs.

Single Covariance:

The screenshot shows a LinkedIn Learning video player. The video content is an Excel spreadsheet titled 'SingleCovariance - Excel'. The spreadsheet has the following data:

	Column1	Column2	Covariance
1	1	3	4.68
2	4	2	1.38
3	7	1	-7.92
4	2	4	0.78
5	3	9	-7.52
6	5	6	0.68
7	8	8	12.58
8	1	4	1.08
9	6	4	-0.42
10	9	2	-10.12
11			-0.48
12			
13			

The spreadsheet also shows a 'Covariance Function' formula: $\frac{\sum(x-\bar{x})(y-\bar{y})}{n}$. The value -0.48 is shown in cell E4.

The video player interface includes a 'Contents' sidebar on the left with the following items:

- Chapter quiz (1 question)
- 5. Measure Covariance and Correlation
 - Visualize what covariance means (2m 38s)
 - Calculate covariance between two columns of data (4m 21s)
 - Calculate covariance among multiple pairs of columns (5m 16s)
 - Visualize what correlation means (4m 42s)
 - Calculate correlation between two columns of data (2m 33s)
 - Calculate correlation among multiple pairs of columns (4m 32s)

The video player also has a search bar for the transcript and a 'Give feedback' button.

Calculate covariance between multiple pairs of columns:

Here for example we took about 4 columns

To find the covariance between

Column 1 and column 1	=COVARIANCE.S(A2:A11,A2:A11)
Column 2 and column 1	=COVARIANCE.S(B2:B11,A2:A11)
Column 3 and column 1	=COVARIANCE.S(C2:C11,A2:A11)
Column 4 and column 1	=COVARIANCE.S(D2:D11,A2:A11)

Multiple Covariance:

The screenshot shows a LinkedIn Learning video player. The video content is an Excel spreadsheet with the following data:

	Column1	Column2	Column3	Column4		Column1	Column2	Column3	Column4
1									
2	1	3	100	70	Column1	8.26667	-0.53333	16.5333	4.22222
3	4	2	104	71	Column2	-0.53333	6.9	-8.17778	1.61111
4	7	1	108	72	Column3	16.5333	-8.17778	60.8444	-1.55556
5	2	4	90	73	Column4	4.22222	1.61111	-1.55556	9.16667
6	3	9	94	74					
7	5	6	103	75					
8	8	8	100	76					
9	1	4	85	77					
10	6	4	98	78					
11	9	2	110	79					
12									
13									

The video player interface includes a table of contents on the left, a search bar, and a transcript area at the bottom. The transcript shows the video title "Calculate covariance among multiple pairs of columns" and a snippet of the instructor's speech: "[Instructor] When you capture data about your business it's likely that you will capture several type of".

Drag the column to get the covariance in the balance columns respectively.

Measure Correlation:

II. Correlation formula =
$$\frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Procedure:

- For each data point, find its deviation from the Mean.
- Multiply the deviation for each pair of data points.
- Find the sum of all these values and divide by the number of terms

Interpreting correlated value:

- 0** - The data that is completely correlated
- $0 > x \geq 1$** - The data that is Positively correlated.
- $-1 \leq x < 0$** - The data that is Negatively correlated.

Calculate Correlation between Two Columns of data:

- The more samples you have, the lower correlation value you need for your result to be significant.
- Take two columns like column 1 and column 2
- To calculate the correlation in excel,using correl function type
 $=\text{correl}(a3:a12,b3:b12)$
- Press enter, we get the correlation values for the column 1 and column 2
- Using correlation two-tailed significant formula table we can find the percent of significant of the correlation values.

Single correlation:

The screenshot shows a LinkedIn Learning video player. The video content is an Excel spreadsheet titled 'SingleCorrelation - Excel'. The spreadsheet has the following data:

	Column1	Column2	Correlation Formula
1			
2	1	3	0.522150671
3	4	5	
4	7	8	
5	2	4	
6	3	9	
7	5	6	
8	8	8	
9	1	4	
10	6	4	
11	9	6	
12			

The video player interface includes a 'Contents' sidebar on the left with the following items:

- Visualize what covariance means (2m 38s)
- Calculate covariance between two columns of data (4m 21s)
- Calculate covariance among multiple pairs of columns (5m 16s)
- Visualize what correlation means (4m 42s)
- Calculate correlation between two columns of data (2m 33s)
- Calculate correlation among multiple pairs of columns (4m 32s)
- Chapter Quiz (3 questions)
- 6. Perform Bayesian Analysis

The video player also features a search bar for the transcript, a 'Give feedback' button, and a system tray at the bottom showing the date and time as 15-03-2021, 10:29.

Two tailed correlation table:

The screenshot shows a LinkedIn Learning video player. The main content is a 'Correlation lookup table (two-tailed)'. The table is as follows:

N	0.1	0.05	0.02	0.01	0.001
5	.80	.88	.93	.96	.99
6	.73	.81	.88	.92	.97
7	.67	.75	.83	.87	.95
8	.62	.71	.79	.83	.93
9	.58	.67	.75	.80	.90
10	.55	.63	.71	.77	.87
15	.44	.51	.59	.64	.76
20	.38	.44	.52	.56	.68
30	.31	.36	.42	.46	.57

The video player interface includes a 'Contents' sidebar on the left with items like 'Visualize what covariance means', 'Calculate covariance between two columns of data', 'Calculate covariance among multiple pairs of columns', 'Visualize what correlation means', 'Calculate correlation between two columns of data', 'Calculate correlation among multiple pairs of columns', and 'Chapter Quiz'. The main video area shows the correlation table and navigation options like 'Overview', 'Q&A', 'Notebook', and 'Transcript'. A search bar is also present.

We took 10 values and 0.52 as correlation, for the first column the value for significance is .55. So what that means is that my value of .52 is almost significant at the 90% level, but in fact, it is not. If I'd had 15 samples, then my value of 0.52 would be significant at both the 90%, or 0.1 and 95%, .51 levels.

Calculate correlation between multiple pairs of columns:

Here for example we took about 4 columns

To find the correlation between

Column 1 and column 1	=CORREL(A2:A11,A2:A11)
Column 2 and column 1	=CORREL(B2:B11,A2:A11)
Column 3 and column 1	=CORREL(C2:C11,A2:A11)
Column 4 and column 1	=CORREL(D2:D11,A2:A11)

And drag the formula for other respective columns.

Multiple correlation:

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D		F	G	H	I	J	K
1	Column1	Column2	Column3	Column4							
2	1	3	100	70							
3	4	2	104	71							
4	7	1	108	72							
5	2	4	90	73							
6	3	9	94	74							
7	5	6	103	75							
8	8	8	100	76							
9	1	4	85	77							
10	6	4	98	78							
11	9	2	110	79							
12											

	Column1	Column2	Column3	Column4
Column1	1	-0.07062	0.737199	0.485032
Column2	-0.07062	1	-0.39912	0.202579
Column3	0.737199	-0.39912	1	-0.06587
Column4	0.485032	0.202579	-0.06587	1

We finally get the correlation values for multiple columns.

Quiz

Chapter Quiz Results

Question 1 of 1

Which random number function generates a real number between 0 and 1?

You are correct!

RAND

Feedback

RAND is used to generate a random number between 0 and 1.

Next chapter

This screenshot shows a quiz result window titled "Chapter Quiz Results". It displays a single question: "Which random number function generates a real number between 0 and 1?". The user has selected "RAND", which is marked as correct with a green checkmark and the text "You are correct!". Below the question, the feedback text states: "RAND is used to generate a random number between 0 and 1." At the bottom right of the window, there is a button labeled "Next chapter".

Chapter Quiz Results

Question 1 of 3

If the covariance value is _____, it means the data sets tend to move in the same direction.

You are correct!

positive

Feedback

A positive covariance means that the data sets move in the same direction.

Next chapter

This screenshot shows a quiz result window titled "Chapter Quiz Results". It displays a question: "If the covariance value is _____, it means the data sets tend to move in the same direction." The user has selected "positive", which is marked as correct with a green checkmark and the text "You are correct!". Below the question, the feedback text states: "A positive covariance means that the data sets move in the same direction." At the bottom right of the window, there is a button labeled "Next chapter".

Search for skills, subjects or software

Chapter Quiz Results

Question 3 of 3

Approximately 68 percent of the values in a normal distribution will be within plus or minus _____.

✔ You are correct!

one standard deviation

Feedback

About 68 percent of the values will be within one standard deviation of the average value.

Next chapter

Search for skills, subjects or software

Chapter Quiz Results

Question 2 of 3

Which formula will return the correlation coefficient between data in cells A1:A5 and B1:B5?

✔ You are correct!

`=CORREL(A1:A5,B1:B5)`

Feedback

This formula correlates the ranges between A1:A5 and B1:B5.

Next chapter

CONCLUSION

The internship was a useful experience. It helped to gain new knowledge and skills. It provides a path to achieve several of our learning goals. I have learned many shortcut formulae in Excel using data.