PROJECT REPORT

ON

CHINA LARGEST COMPANIES

(STATISTICAL ANALYSIS USING DATA ANALYSIS FEATURE OF EXCEL)

INDEX

S.NO	TOPIC	PAGE NO
1.	Introduction	
2.	**Excel Analysis Tool Pak**	
2.1	Anova Test	
2.2	Correlation	
2.3	Covariance	
2.4	Descriptive Statistics	
2.5	F-test: Two Sample for variances	
2.6	Rank & Percentile	
2.7	Regression	
2.8	T-test: pair two sample for means	
2.9	T-test: two sample assuming equal variance	
2.10	T-test: two sample assuming unequal variance	
2.11	Z-Test: two sample for means	
	Analysis Practical Approach	
3.	Dashboard	

INTRODUCTION

Data analysis is a process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, and is used in different business, science, and social science domains. In today's business world, data analysis plays a role in making decisions more scientific and helping businesses operate more effectively.

Data mining is a particular data analysis technique that focuses on statistical modelling and knowledge discovery for predictive rather than purely descriptive purposes, while business intelligence covers data analysis that relies heavily on aggregation, focusing mainly on business information. In statistical applications, data analysis can be divided into descriptive statistics, exploratory data analysis (EDA), and confirmatory data analysis (CDA). EDA focuses on discovering new features in the data while CDA focuses on confirming or falsifying existing hypotheses. Predictive analytics focuses on the application of statistical models for predictive forecasting or classification, while text analytics applies statistical, linguistic, and structural techniques to extract and classify information from textual sources, a species of unstructured data. All of the above are varieties of data analysis. Data integration is a precursor to data analysis, and data analysis is closely linked to data visualization and data dissemination.

➢ Excel Analysis Tool Pak:

The Analysis Tool Pak includes the tools described in the following sections. To access these tools, Data Analysis in the Analysis group on the Data tab. If the Data Analysis command is not available, we need to load the Analysis Tool Pak add-in program.

ANOVA TEST:

The Anova analysis tools provide different types of variance analysis. The tool that we should use depends on the number of factors and the number of samples that we have from the Sample that we want to test.

1.1 Anova: Single Factor

This tool performs a simple analysis of variance on data for two or more samples. The analysis provides a test of the hypothesis that each sample is drawn from the same underlying probability distribution against the alternative hypothesis that underlying probability distributions are not the same for all samples. If there are only two samples, we can use the worksheet function **T.TEST**. With more than two samples, there is no convenient generalization of **T.TEST**, and the Single Factor Anova model can be called upon instead.

1.2 Anova: Two-Factor with Replication

This analysis tool is useful when data can be classified along two different dimensions. For example, in an experiment to measure the height of plants, the plants may be given different brands of fertilizer (for example, A, B, C) and might also be kept at different temperatures (for example, low, high). For each of the six possible pairs of {fertilizer, temperature}, we have an equal number of observations of plant height. Using this Anova tool, we can test:

- Whether the heights of plants for the different fertilizer brands are drawn from the same underlying population. Temperatures are ignored for this analysis.

- Whether the heights of plants for the different temperature levels are drawn from the same underlying population. Fertilizer brands are ignored for this analysis.

Whether having accounted for the effects of differences between fertilizer brands found in the first bulleted point and differences in temperatures found in the second bulleted point, the six samples representing all pairs of {fertilizer, temperature} values are drawn from the same population. The alternative hypothesis is that there are effects due to specific {fertilizer, temperature} pairs over and above the differences that are based on fertilizer alone or on temperature alone.

1.3 Anova: Two-Factor without Replication

This analysis tool is useful when data is classified on two different dimensions as in the Two-Factor case With Replication. However, for this tool it is assumed that there is only a single observation for each pair (for example, each {fertilizer, temperature} pair in the preceding example).

Correlation

The CORREL and PEARSON worksheet functions both calculate the correlation coefficient between two measurement variables when measurements on each variable are observed for each of N subjects. (Any missing observation for any subject causes that subject to be ignored in the analysis.) The Correlation analysis tool is particularly useful when there are more than two measurement variables for each of N subjects. It provides an output table, a correlation matrix that shows the value of CORREL (or PEARSON) applied to each possible pair of measurement variables.

The correlation coefficient, like the covariance, is a measure of the extent to which two measurement variables "vary together." Unlike the covariance, the correlation coefficient is scaled so that its value is independent of the units in which the two measurement variables are expressed. (For example, if the two measurement variables are weight and height, the value of the correlation coefficient is unchanged if weight is converted from pounds to kilograms.) The value of any correlation coefficient must be between -1 and +1 inclusive.

We can use the correlation analysis tool to examine each pair of measurement variables to determine whether the two measurement variables tend to move together — that is, whether large values of one variable tend to be associated with large values of the other (positive correlation), whether small values of one variable tend to be associated with large values of the other (negative correlation), or whether values of both variables tend to be unrelated (correlation near 0 (zero)).

Covariance

1. The Correlation and Covariance tools can both be used in the same setting, when we have N different measurement variables observed on a set of individuals. The Correlation and Covariance tools each give an output table, a matrix that shows the correlation coefficient or covariance, respectively, between each pair of measurement variables. The difference is that correlation coefficients are scaled to lie between -1 and +1 inclusive. Corresponding covariances are not scaled. Both the correlation coefficient and the covariance are measures of the extent to which two variables "vary together."
2. The Covariance tool computes the value of the worksheet function COVARIANCE.P for each pair of measurement variables. (Direct use of COVARIANCE.P rather than the Covariance tool is a reasonable alternative when there are only two measurement variables, that is, N=2.) The entry on the diagonal of the Covariance tool's output table in row i, column i is the covariance of the i-th measurement variable with itself. This is just the population variance for that variable, as calculated by the worksheet function VAR.P.
3. We can use the Covariance tool to examine each pair of measurement variables to determine whether the two measurement variables tend to move together — that is, whether large values of one variable tend to be associated with large values of the other (positive covariance), whether small values of one variable tend to be associated with large values of the other (negative covariance), or whether values of both variables tend to be unrelated (covariance near 0 (zero)).

Descriptive Statistics:

The Descriptive Statistics analysis tool generates a report of univariate statistics for data in the input range, providing information about the central tendency and variability of wer data.

F-TEST: Two Sample for Variances

1. The F-Test Two-Sample for Variances analysis tool performs a two-sample F-test to compare two population variances.
2. For example, we can use the F-Test tool on samples of times in a swim meet for each of two teams. The tool provides the result of a test of the null hypothesis that these two samples come from distributions with equal variances, against the alternative that the variances are not equal in the underlying distributions.
3. The tool calculates the value f of an F-statistic (or F-ratio). A value of f close to 1 provides evidence that the underlying population variances are equal. In the output table, if f < 1 "P(F <= f) one-tail" gives the probability of observing a value of the F-statistic less than f when population variances are equal, and "F Critical one-tail" gives the critical value less than 1 for the chosen significance level, Alpha. If f > 1, "P(F <= f) one-tail" gives the probability of observing a value of the F-statistic greater than f when population variances are equal, and "F Critical one-tail" gives the critical value greater than 1 for Alpha.

.

Rank and percentile:

The Rank and Percentile analysis tool produces a table that contains the ordinal and percentage rank of each value in a data set. We can analyze the relative standing of values in a data set. This tool uses the worksheet functions RANK.EQ andPERCENTRANK.INC. If we want to account for tied values, use the RANK.EQ function, which treats tied values as having the same rank, or use the RANK.AVG function, which returns the average rank for the tied values.

Regression:

The Regression analysis tool performs linear regression analysis by using the "least squares" method to fit a line through a set of observations. We can analyze how a single dependent variable is affected by the values of one or more independent variables. For example, we can analyze how an athlete's performance is affected by such factors as age, height, and weight. We can apportion shares in the performance measure to each of these three factors, based on a set of performance data, and then use the results to predict the performance of a new, untested athlete.

T-test:

The Two-Sample t-Test analysis tools test for equality of the population means that underlie each sample. The three tools employ different assumptions: that the population variances are equal, that the population variances are not equal, and that the two samples represent before-treatment and after-treatment observations on the same subjects.

For all three tools below, a t-Statistic value, t, is computed and shown as "t Stat" in the output tables. Depending on the data, this value, t, can be negative or nonnegative. Under the assumption of equal underlying population means, if t < 0, "P (T <= t) one-tail" gives the probability that a value of the t-Statistic would be observed that is more negative than t. If t >=0, "P(T <= t) one-tail" gives the probability that a value of the t-Statistic would be observed that is more positive than t. "t Critical one-tail" gives the cutoff value, so that the probability of observing a value of the t-Statistic greater than or equal to "t Critical one-tail" is Alpha.

"P(T <= t) two-tail" gives the probability that a value of the t-Statistic would be observed that is larger in absolute value than t. "P Critical two-tail" gives the cutoff value, so that the probability of an observed t-Statistic larger in absolute value than "P Critical two-tail" is Alpha.

1. t-Test: Paired Two Sample For Means

We can use a paired test when there is a natural pairing of observations in the samples, such as when a sample group is tested twice — before and after an experiment. This analysis tool and its formula perform a paired two-sample Student's t-Test to determine whether observations that are taken before a treatment and observations taken after a treatment are likely to have come from distributions with equal population means. This t-Test form does not assume that the variances of both populations are equal.

2. t-Test: Two-Sample Assuming Equal Variances

This analysis tool performs a two-sample student's t-Test. This t-Test form assumes that the two data sets came from distributions with the same variances. It is referred to as a homoscedastic t-Test. We can use this t-Test to determine whether the two samples are likely to have come from distributions with equal population means.

3. t-Test: Two-Sample Assuming Unequal Variances

This analysis tool performs a two-sample student's t-Test. This t-Test form assumes that the two data sets came from distributions with unequal variances. It is referred to as a heteroscedastic t-Test. As with the preceding Equal Variances case, we can use this t-Test to determine whether the two samples are likely to have come from distributions with equal population means. Use this test when there are distinct subjects in the two samples. Use the Paired test, described in the follow example, when there is a single set of subjects and the two samples represent measurements for each subject before and after a treatment.

Z-Test:

The z-Test: Two Sample for Means analysis tool performs a two sample z-Test for means with known variances. This tool is used to test the null hypothesis that there is no difference between two population means against either one-sided or two-sided alternative hypotheses. If variances are not known, the worksheet function Z.TEST should be used instead.

When we use the z-Test tool, be careful to understand the output. "P(Z <= z) one-tail" is really P(Z >= ABS(z)), the probability of a z-value further from 0 in the same direction as the observed z value when there is no difference between the population means. "P (Z <= z) two-tail" is really P(Z >= ABS(z) or Z <= -ABS(z)), the probability of a z-value further from 0 in either direction than the observed z-value when there is no difference between the population means. The two-tailed result is just the one-tailed result multiplied by 2. The z-Test tool can also be used for the case where the null hypothesis is that there is a specific nonzero value for the difference between the two population means. For example, we can use this test to determine differences between the performances of two car models.

DATA ANALYSIS OF CHINA LARGEST COMPANIES

- ANNOVA TEST: SINGLE FACTOR
- ANNOVA TEST: TWO FACTOR WITHOUT REPLICATION
- ANNOVA TEST: TWO FACTOR WITH REPLICATION
- CORRELATION
- COVARIANCE
- DESCRIPTIVE STATISTICS
- F-TEST: TWO SAMPLE FOR VARIANCE
- RANK AND PERCENTILE
- Z-TEST:TWO SAMPLE FOR MEAN
- REGRESSION
- T-TEST: PAIRED
- T-TEST: TWO SAMPLE - EQUAL VARIANCES
- T-TEST: TWO SAMPLE -UNEQUAL VARIANCES
- DATASET

Anova: Single Factor

BACK TO DASHBOARD

SUMMARY				
Groups	Count	Sum	Average	Variance
Sales ($billion)	135	3024.5	22.4037037	2288.327523
Profits ($billion)	135	275.4	2.04	27.60002985
Assets ($billion)	135	17198.9	127.3992593	164794.8532
Market Value ($billion)	135	2762.8	20.46518519	1540.128555

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1313949.347	3	437983.1156	10.38792183	1.2E-06	2.621533
Within Groups	22599221.85	536	42162.72733			
Total	23913171.2	539				

Conclusion:

P value is greater than alpha value accept null hyphothesis and do not reject

F value is greater than F crit value then

reject null Hypothesis (F value > F crit Value)

Anova: Two-Factor Without Replication

SUMMARY	Count	Sum	Average	Variance
ICBC	4	3223.4	805.85	1798039
China Construction Bank	4	2586.7	646.675	1134620
Agricultural Bank of China	4	2401	600.25	1034968
PetroChina	4	936.2	234.05	21942.26
Bank of China	4	2285.7	571.425	952564.4
Sinopec-China Petroleum	4	728.7	182.175	29425.2
Bank of Communications	4	956	239	164368.2
Ping An Insurance Group	4	567.5	141.875	44491.61
China Merchants Bank	4	626.8	156.7	67931.37
China Life Insurance	4	449.5	112.375	17550.03
China Minsheng Banking	4	593.4	148.35	60150.73
China Shenhua Energy	4	188.4	47.1	900.8067
Shanghai Pudong Development	4	566.9	141.725	58616.41
China Citic Bank	4	536.5	134.125	51685.88
China Telecom	4	176.7	44.175	1206.269
Industrial Bank	4	443.3	110.825	32950.1
SAIC Motor	4	153.1	38.275	934.1625
China State Construction	4	175.7	43.925	1613.243
People's Insurance Company	4	177.3	44.325	2217.896
China Everbright Bank	4	312.5	78.125	17299.2
China Pacific Insurance	4	166	41.5	2207.127
China Communications Construction	4	130.1	32.525	957.0558
Ping An Bank	4	291	72.75	15182.34
China Vanke	4	96.7	24.175	633.2758
China Railway Construction	4	163.3	40.825	1645.929
China Railway Group	4	176.5	44.125	2004.956
Baoshan Iron & Steel	4	86.7	21.675	300.6092
Huaneng Power International	4	78.1	19.525	280.1892
Huaxia Bank	4	220.7	55.175	8985.156
China Coal Energy	4	56.8	14.2	133.2467
Bank of Beijing	4	172.5	43.125	5262.363
Dongfeng Motor Group	4	52.1	13.025	73.10917
Poly Real Estate	4	63.4	15.85	287.2967
New China Life Insurance	4	109.5	27.375	1246.189
Tencent Holdings	4	86.1	21.525	857.0358
China National Building	4	53.4	13.35	169.35
Daqin Railway	4	42.7	10.675	57.10917
Datang International Power	4	63.6	15.9	335.4267
Evergrande Real Estate	4	49.8	12.45	167.7167
Jiangxi Copper	4	49.1	12.275	98.60917

BACK TO DASHBOARD

China Shipbuilding Industry	4	47.3	11.825	104.1025
Gree Electric Appliances	4	41.1	10.275	39.90917
Yanzhou Coal Mining	4	40	10	49.62667
Lenovo Group	4	56	14	146.94
CSR	4	41.4	10.35	50.31
Anhui Conch Cement	4	38.1	9.525	46.24917
China Yangtze Power	4	49.1	12.275	140.3425
Sinohydro Group	4	49.9	12.475	132.5825
Zoomlion Heavy Industry	4	33.1	8.275	28.6625
Country Garden Holdings	4	37.8	9.45	78.13667
Citic Securities	4	53	13.25	188.5633
GD Power Development	4	45.8	11.45	152.1967
Metallurgical Corp of China	4	95.1	23.775	698.7558
Shanghai Electric Group	4	38.8	9.7	56.36667
Sany Heavy Industry	4	30.6	7.65	22.69667
Aluminium Corp of China	4	59.5	14.875	190.8558
China Southern Airlines	4	42.4	10.6	65.36667
China CNR	4	37.3	9.325	47.68917
China Eastern Airlines	4	39.1	9.775	69.71583
Zijin Mining Group	4	30	7.5	22.22667
Suning Appliance	4	32.5	8.125	33.9025
Kweichow Moutal	4	41.7	10.425	166.2425
Sinopharm Group	4	42.6	10.65	78.83
Great Wall Motor	4	28.6	7.15	30.20333
Baidu	4	42.2	10.55	168.4367
Longfor Properties	4	30.8	7.7	46.46667
Haitong Securities	4	39.1	9.775	103.8625
Chonqing Rural Bank	4	79	19.75	1107.503
Tingyi Holding	4	32.2	8.05	35.64333
Sun Art Retail Group	4	33.8	8.45	36.96333
Fosun International	4	38.9	9.725	127.3692
Shanghai International Port	4	28.4	7.1	30.72667
China Cosco Holdings	4	45.4	11.35	142.0567
Shimao Property Holdings	4	30.5	7.625	63.9825
China Oilfield Services	4	27.3	6.825	31.20917
Wuliangye Yibin	4	24.6	6.15	36.53667
Weichai Power	4	25.7	6.425	18.0825
Henan Shuanghui Investment	4	22.1	5.525	27.9625
ZTE	4	35.8	8.95	55.95
Hebei Iron & Steel	4	48.3	12.075	134.3025
Inner Mongolia Yitai	4	21.1	5.275	10.44917
Bank of Ningbo	4	49.1	12.275	377.6692
Greentown China Holdings	4	26.9	6.725	53.5425

Guangzhou R&F	4	26.3	6.575	41.2025
Qingdao Haier	4	26.5	6.625	25.53583
Want Want China	4	27.3	6.825	76.62917
Bank of Nanjing	4	51.8	12.95	450.5767
Sino-Ocean Land Holdings	4	29	7.25	77.23
Huadian Power International	4	40.8	10.2	130.7467
China Merchants Property	4	27.4	6.85	55.63667
Jiangsu Yanghe Brewery	4	17.9	4.475	28.80917
China Longyuan Power	4	27.6	6.9	56.04667
GF Securities	4	27.4	6.85	51.81667
Wuhan Iron & Steel	4	36.3	9.075	63.4425
Agile Property Holdings	4	24.3	6.075	36.04917
Soho China	4	20.8	5.2	25.24667
Inner Mongolia Rare-Earth	4	16.4	4.1	26.81333
Shenzhen Overseas	4	22	5.5	23.38
Shanghai Construction	4	30.8	7.7	52.45333
BYD	4	26.7	6.675	22.1425
Angang Steel	4	31.9	7.975	58.87583
Shanxi Taigang Stainless	4	29.2	7.3	46.96
Yang Quan Coal Industry	4	18.2	4.55	9.956667
Hengan International Group	4	18.6	4.65	25.19
Shanghai Pharmaceuticals	4	24.8	6.2	19.98
China National Chemical	4	21.2	5.3	10.71333
New Hope Liuhe	4	18.9	4.725	22.00917
BBMG	4	23.3	5.825	28.92917
Huatai Securities	4	24.7	6.175	41.48917
Xiamen C&D	4	23.7	5.925	32.03583
China International Marine	4	24.1	6.025	19.64917
China Merchants Securities	4	22	5.5	31.06667
Hunan Valin Steel	4	26.3	6.575	49.97583
China Communications Services	4	22.1	5.525	16.00917
Dongfang Electric	4	23.6	5.9	25.9
Inner Mongolia Yili	4	18.4	4.6	13.63333
Maanshan Iron & Steel	4	25.7	6.425	42.86917
TCL Corp	4	27.3	6.825	36.69583
Shandong Gold-Mining	4	16.2	4.05	11.97667
Netease	4	11.8	2.95	7.696667
Jizhong Energy Resources	4	17.4	4.35	6.856667
Minmetals Development	4	36	9	133.6733
Gemdale	4	22.9	5.725	35.5225
China Gezhouba	4	23.7	5.925	29.4025
Great Wall Technology	4	23.9	5.975	53.50917
Luzhou Laojiao	4	10.3	2.575	7.529167

Hainan Airlines	4	24.1	6.025	38.83583
China Hongqiao Group	4	14.9	3.725	6.6425
Zhongjin Gold	4	15.5	3.875	8.2425
Shanxi Lu'an Environmental	4	16.6	4.15	8.91
Qinghai Salt Lake	4	27.1	6.775	38.6425
Inner Mongolia Baotou Steel	4	21.2	5.3	12.42
Shanghai Material Trading	4	14.6	3.65	37.63
Tongling Nonferrous Metals	4	22.1	5.525	25.58917
Hikvision	4	14.6	3.65	33.45667
Sales ($billion)	135	3024.5	22.4037	2288.328
Profits ($billion)	135	275.4	2.04	27.60003
Assets ($billion)	135	17198.9	127.3993	164794.9
Market Value ($billion)	135	2762.8	20.46519	1540.129

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	7287532	134	54384.57	1.427837	0.004434	1.25215
Columns	1313949	3	437983.1	11.49901	3.01E-07	2.627103
Error	15311690	402	38088.78			
Total	23913171	539				

conclusion:

If the F-value (1.42)larger than the f critical value (1.25). so, there is a statistical significance to the results,
leading to reject the null hypothesis.

If the p-value(0.004) smaller than chosen alpha(0.05) level. so, that would also lead to reject the null hypothesis.

Anova: Two-Factor With Replication

SUMMARY	Sales ($billion)	Profits ($billion)	Assets ($billion)	Market Value ($billion)	Total
ICBC					
Count	67	67	67	67	268
Sum	2519.5	249.4	16364.7	2322.7	21456.3
Average	37.60447761	3.72238806	244.2492537	34.66716418	80.0608209
Variance	4152.984071	50.1487336	306945.798	2705.457693	86778.42951
Chonqing Rural Bank					
Count	67	67	67	67	268
Sum	504.2	25.8	832.9	427.8	1790.7
Average	7.525373134	0.385074627	12.43134328	6.385074627	6.681716418
Variance	26.65949796	0.182804161	118.0797603	14.45431931	57.84164948
Total					
Count	134	134	134	134	
Sum	3023.7	275.2	17197.6	2750.5	
Average	22.56492537	2.053731343	128.3402985	20.5261194	
Variance	2301.997633	27.78190327	165913.4635	1551.203448	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Sample	721522.0585	1	721522.0585	18.38192179	2.14938E-05	3.859131
Columns	1324145.078	3	441381.6928	11.2449005	3.66903E-07	2.621786
Interaction	1136230.839	3	378743.6131	9.649095816	3.28434E-06	2.621786
Within	20724908.48	528	39251.72061			
Total	23906806.46	535				

Conclusion:

we are going to reject the null hypothesis or not, we'll depend on two factors:

If the F-value (f)is larger than the f critical value (f crit)

If the p-value is smaller than your chosen alpha level.

accept null hypothesis

BACK TO DASHBOARD

Correlation between Sales vs Profits vs Assets vs Market value

	Sales ($billion)	Profits ($billion)	Assets ($billion)	Market Value ($billion)
Sales ($billion)	1			
Profits ($billion)	0.590117708	1		
Assets ($billion)	0.409830163	0.945901201	1	
Market Value ($billion)	0.725870428	0.923257086	0.80134643	1

Conclusion:

Assets and Profits of the China largest companies is more correlated

BACK TO DASHBOARD

Covariance between Sales vs Profits vs Assets vs Market value

	Sales ($billion)	Profits ($billion)	Assets ($billion)	Market Value ($billion)
Sales ($billion)	2271.376949			
Profits ($billion)	147.2053333	27.39558519		
Assets ($billion)	7899.613855	2002.365733	163574.1506	
Market Value ($billion)	1352.594796	188.9413185	12671.89731	1528.720195

Conclusion				
Market value and Assets values of china largest companies is more related				

BACK TO DASHBOARD

DESCRIPTIVE STATISTICS

Sales ($billion)			Profits ($billion)			Assets ($billion)			Market Value ($billion)	
Mean	22.4037037		Mean	2.04		Mean	127.3992593		Mean	20.46519
Standard Error	4.11710751		Standard Error	0.452155577		Standard Error	34.93855487		Standard Error	3.377626
Median	10.3		Median	0.6		Median	15.5		Median	8.7
Mode	6.3		Mode	0.5		Mode	7.2		Mode	11.8
Standard Deviation	47.83646646		Standard Deviation	5.253573056		Standard Deviation	405.9493234		Standard Deviation	39.24447
Sample Variance	2288.327523		Sample Variance	27.60002985		Sample Variance	164794.8532		Sample Variance	1540.129
Kurtosis	41.72033759		Kurtosis	24.96226909		Kurtosis	26.1049189		Kurtosis	20.56941
Skewness	5.950324843		Skewness	4.778898908		Skewness	5.008427277		Skewness	4.344175
Range	410.9		Range	39.3		Range	2812.2		Range	261
Minimum	0.8		Minimum	-1.5		Minimum	1.3		Minimum	0.2
Maximum	411.7		Maximum	37.8		Maximum	2813.5		Maximum	261.2
Sum	3024.5		Sum	275.4		Sum	17198.9		Sum	2762.8
Count	135		Count	135		Count	135		Count	135
Largest(1)	411.7		Largest(1)	37.8		Largest(1)	2813.5		Largest(1)	261.2
Smallest(1)	0.8		Smallest(1)	-1.5		Smallest(1)	1.3		Smallest(1)	0.2

BACK TO DASHBOARD

F-Test Two-Sample for Variances

	Profits ($billion)	Sales ($billion)
Mean	2.04	22.404
Variance	27.600	2288.328
Observations	135	135
df	134	134
F	0.012	
P(F<=f) one-tail	0	
F Critical one-tail	0.752	

Conclusion:

if F < F Critical one-tail, we accept the null hypothesis. This is the case, 0.012 <0.752.

Therefore, we accept the null hypothesis. The variances of the two Samples are equal.

BACK TO DASHBOARD

Rank and Percentile

Point	Sales ($billion)	Rank	Percent
6	411.7	1	100.00%
4	308.9	2	99.20%
1	134.8	3	98.50%
2	113.1	4	97.70%
3	103	5	97.00%
5	98.1	6	96.20%
18	76.7	7	94.70%
26	76.7	7	94.70%
17	75	9	94.00%
25	74.5	10	93.20%
10	63.2	11	92.50%
8	51.1	12	91.70%
22	45.7	13	91.00%
15	44.9	14	90.20%
7	43.5	15	89.50%
19	41.3	16	88.80%
12	39.7	17	88.00%
27	35.4	18	87.30%
53	34.1	19	86.50%
44	29.6	20	85.80%
9	28.4	21	85.00%
11	27.9	22	84.30%
21	27.3	23	83.50%
122	25.7	24	82.80%
13	25.4	25	82.00%
40	25	26	81.30%
14	24.4	27	80.50%
56	23.7	28	79.80%
63	21.5	29	79.10%
80	21.2	30	78.30%
28	21.1	31	77.60%
32	19.7	32	76.80%
16	18.7	33	76.10%
34	18	34	74.60%
48	18	34	74.60%

94	16.1	36	73.80%
57	15.7	37	73.10%
24	15.5	38	72.30%
102	15.3	39	71.60%
125	15.2	40	70.10%
131	15.2	40	70.10%
61	14.9	42	69.40%
99	14.5	43	68.60%
58	14.2	44	67.90%
45	14.1	45	67.10%
73	14	46	66.40%
36	13.8	47	65.60%
30	13.6	48	64.90%
20	13.5	49	63.40%
59	13.5	49	63.40%
42	13.3	51	61.90%
79	13.3	51	61.90%
23	13.2	53	61.10%
110	12.8	54	59.70%
133	12.8	54	59.70%
85	12.6	56	58.90%
70	12.3	57	57.40%
101	12.3	57	57.40%
38	12.2	59	55.90%
134	12.2	59	55.90%
54	12.1	61	55.20%
117	11.8	62	54.40%
113	11.7	63	53.70%
107	11.4	64	52.90%
118	11	65	52.20%
105	10.8	66	51.40%
29	10.5	67	50.70%
39	10.3	68	50.00%
114	9.8	69	49.20%
33	9.6	70	48.50%
89	9.4	71	47.70%
41	9.2	72	45.50%
43	9.2	72	45.50%

69	9.2	72	45.50%
111	8.6	75	44.70%
124	8.3	76	44.00%
71	8.2	77	43.20%
55	8.1	78	42.50%
52	8	79	41.00%
103	8	79	41.00%
49	7.6	81	38.80%
60	7.6	81	38.80%
77	7.6	81	38.80%
100	7.4	84	38.00%
37	7.2	85	36.50%
46	7.2	85	36.50%
35	7	87	35.80%
106	6.9	88	35.00%
132	6.8	89	34.30%
50	6.6	90	32.80%
64	6.6	90	32.80%
31	6.3	92	30.50%
78	6.3	92	30.50%
119	6.3	92	30.50%
115	6	95	28.30%
116	6	95	28.30%
121	6	95	28.30%
83	5.6	98	27.60%
129	5.3	99	26.80%
108	5.1	100	26.10%
81	5	101	25.30%
84	4.8	102	23.80%
95	4.8	102	23.80%
88	4.6	104	23.10%
72	4.5	105	21.60%
74	4.5	105	21.60%
66	4.4	107	20.10%
127	4.4	107	20.10%
128	3.9	109	19.40%
62	3.8	110	17.90%
123	3.8	110	17.90%

68	3.6	112	17.10%
65	3.5	113	15.60%
75	3.5	113	15.60%
86	3.4	115	14.90%
47	3.3	116	13.40%
90	3.3	116	13.40%
130	3.1	118	12.60%
76	3	119	11.10%
98	3	119	11.10%
92	2.7	121	10.40%
96	2.5	122	9.70%
82	2.4	123	8.20%
104	2.4	123	8.20%
51	2.2	125	6.70%
87	2.2	125	6.70%
91	2	127	5.90%
97	1.8	128	5.20%
67	1.6	129	4.40%
109	1.3	130	2.20%
120	1.3	130	2.20%
126	1.3	130	2.20%
112	1.1	133	1.40%
93	1	134	0.70%
135	0.8	135	0.00%

Conclusion:

We can see the output above from the tool. Excel has ranked every sales and given them their exact percentile (the Percent column). However, rather than using the Company names, there is a column called Point, which is the row number of each company.

BACK TO DASHBOARD

z-Test: Two Sample for Means

	Sales ($billion)	Profits ($billion)
Mean	22.40	2.04
Known Variance	2288.33	27.60002985
Observations	135	135
Hypothesized Mean Difference	0	
z	4.9166	
P(Z<=z) one-tail	4.40395E-07	
z Critical one-tail	1.644853627	
P(Z<=z) two-tail	8.80791E-07	
z Critical two-tail	1.959963985	

BACK TO DASHBOARD

Conclusion:

Cells B4 and C4 contain the mean of each sample, Variable 1 = Sales and Variable 2 = Profit.

Cells B5 and B5 contain our entries for that variance of each sample.

Cells B6and C6 contain the number of observations in each sample.

Cell B7 contains our entry for the Hypothesized Mean Difference.

Cell B8 contains the result of the actual z-test. We will compare this value to the z-Critical two-tail statistic.

Note: Use a one-tail test if you have a direction in your hypothesis, i.e. if testing that a value is above or below some level.

In this example P(Z <= z) two tail (1) gives us the probability that a value of the z-Statistic (4.9166) would be observed that is larger in absolute value than z Critical two-tail (1.96).

Since the p-value is larger than our Alpha (0.05), we cannot reject the null hypothesis that there is no significant difference in the means of each sample.

Regression

BACK TO DASHBOARD

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.923257086
R Square	0.852403646
Adjusted R Square	0.851293899
Standard Error	2.025905431
Observations	135

ANOVA					
	df	SS	MS	F	Significance F
Regression	1	3152.533056	3152.533056	768.1062729	4.13856E-57
Residual	133	545.8709442	4.104292813		
Total	134	3698.404			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-0.48938313	0.196803094	-2.486663801	0.0141329	-0.878652027	-0.100114233	-0.878652027	-0.100114233
Market Value ($billion)	0.123594441	0.004459522	27.71473025	4.13856E-57	0.11477368	0.132415203	0.11477368	0.132415203

t-Test: Paired Two Sample for Means

	Sales ($billion)	Profits ($billion)
Mean	22.4037037	2.04
Variance	2288.327523	27.60002985
Observations	135	135
Pearson Correlation	0.590117708	
Hypothesized Mean Difference	0	
df	134	
t Stat	5.26527553	
P(T<=t) one-tail	2.71349E-07	
t Critical one-tail	1.656304542	
P(T<=t) two-tail	5.42698E-07	
t Critical two-tail	1.977825758	

BACK TO DASHBOARD

Conclusion:			
P value-	Paired two san	5.14E-07	Fail to reject null hypothesis ==> accept Null Hypothesis

t-Test: Two-Sample Assuming Equal Variances

	Sales ($billion)	Profits ($billion)
Mean	22.4037037	2.04
Variance	2288.327523	27.60002985
Observations	135	135
Pooled Variance	1157.963777	
Hypothesized Mean Difference	0	
df	268	
t Stat	4.916558021	
P(T<=t) one-tail	7.68556E-07	
t Critical one-tail	1.650559157	
P(T<=t) two-tail	1.53711E-06	
t Critical two-tail	1.968855173	

BACK TO DASHBOARD

Conclusion:			
P value-	T test with equa	1.53711	Fail to reject null hypothesis ==> accept Null Hypothesis

t-Test: Two-Sample Assuming Unequal Variances

	Sales ($billion)	Profits ($billion)
Mean	22.4037037	2.04
Variance	2288.327523	27.60002985
Observations	135	135
Hypothesized Mean Difference	0	
df	137	
t Stat	4.916558021	
P(T<=t) one-tail	1.2401E-06	
t Critical one-tail	1.65605208	
P(T<=t) two-tail	2.4802E-06	
t Critical two-tail	1.977431212	

BACK TO DASHBOARD

P value-	Paired two sam	5.42698E-07	Fail to reject null hypothesis ==> accept Null Hypothesis
P value-	T test with equa	1.53711E-06	Fail to reject null hypothesis ==> accept Null Hypothesis
P value-	T test without u	2.47629E-06	Fail to reject null hypothesis ==> accept Null Hypothesis

Global Rank	Company	Sales ($billion)	Profits ($billion)	Assets ($billion)	Market Value ($billion)
1	ICBC	134.8	37.8	2813.5	237.3
2	China Construction Bank	113.1	30.6	2241	202
8	Agricultural Bank of China	103	23	2124.2	150.8
9	PetroChina	308.9	18.3	347.8	261.2
11	Bank of China	98.1	22.1	2033.8	131.7
26	Sinopec-China Petroleum	411.7	10.1	200	106.9
54	Bank of Communications	43.5	9.4	846.4	56.7
83	Ping An Insurance Group	51.1	3.2	456.2	57
101	China Merchants Bank	28.4	7.3	547	44.1
106	China Life Insurance	63.2	1.8	304.6	79.9
107	China Minsheng Banking	27.9	6.1	515.5	43.9
115	China Shenhua Energy	39.7	7.7	70.2	70.8
125	Shanghai Pudong Development	25.4	5.4	504.5	31.6
128	China Citic Bank	24.4	4.9	474.7	32.5
139	China Telecom	44.9	2.4	87.4	42
142	Industrial Bank	18.7	4.1	382.3	38.2
167	SAIC Motor	75	3.3	48.1	26.7
206	China State Construction	76.7	2.2	79.9	16.9
226	People's Insurance Company	41.3	1.1	110.5	24.4
240	China Everbright Bank	13.5	2.9	275.1	21
278	China Pacific Insurance	27.3	0.8	109.3	28.6
280	China Communications Construction	45.7	1.9	69.6	12.9
283	Ping An Bank	13.2	2.2	257.3	18.3
314	China Vanke	15.5	2	60.3	18.9
316	China Railway Construction	74.5	1.3	77.1	10.4
322	China Railway Group	76.7	1.2	88.4	10.2
370	Baoshan Iron & Steel	35.4	1.2	36.7	13.4
404	Huaneng Power International	21.1	0.9	41.2	14.9
407	Huaxia Bank	10.5	1.5	197.2	11.5
473	China Coal Energy	13.6	1.5	29.5	12.2
488	Bank of Beijing	6.3	1.4	151.7	13.1
503	Dongfeng Motor Group	19.7	1.4	19.2	11.8
507	Poly Real Estate	9.6	1.3	40.3	12.2
508	New China Life Insurance	18	0.5	79.2	11.8
591	Tencent Holdings	7	2	12.1	65
628	China National Building	13.8	0.9	31.2	7.5
635	Daqin Railway	7.2	1.9	14.9	18.7
639	Datang International Power	12.2	0.6	42.4	8.4
648	Evergrande Real Estate	10.3	1.5	31.1	6.9
650	Jiangxi Copper	25	0.8	12.5	10.8
653	China Shipbuilding Industry	9.2	0.7	25.3	12.1
668	Gree Electric Appliances	13.3	0.8	13.5	13.5

674	Yanzhou Coal Mining	9.2	1	18	11.8
692	Lenovo Group	29.6	0.5	15.5	10.4
710	CSR	14.1	0.6	16.8	9.9
728	Anhui Conch Cement	7.2	1	14	15.9
740	China Yangtze Power	3.3	1.2	25.2	19.4
774	Sinohydro Group	18	0.6	25.8	5.5
779	Zoomlion Heavy Industry	7.6	1.2	13.8	10.5
783	Country Garden Holdings	6.6	1.1	21.9	8.2
798	Citic Securities	2.2	0.7	27	23.1
799	GD Power Development	8	0.6	29.2	8
830	Metallurgical Corp of China	34.1	-1.1	56.2	5.9
848	Shanghai Electric Group	12.1	0.4	18.3	8
851	Sany Heavy Industry	8.1	1.4	8.1	13
853	Aluminium Corp of China	23.7	-1.3	28.7	8.4
859	China Southern Airlines	15.7	0.4	18.3	8
863	China CNR	14.2	0.5	15.4	7.2
868	China Eastern Airlines	13.5	0.5	19.4	5.7
880	Zijin Mining Group	7.6	0.8	10.8	10.8
896	Suning Appliance	14.9	0.8	9.4	7.4
934	Kweichow Moutal	3.8	2.1	6.2	29.6
952	Sinopharm Group	21.5	0.3	12.8	8
957	Great Wall Motor	6.6	0.9	6.8	14.3
960	Baidu	3.5	1.7	7.3	29.7
965	Longfor Properties	4.4	1	16.8	8.6
969	Haitong Securities	1.6	0.5	20.3	16.7
991	Chonqing Rural Bank	3.6	0.8	69.6	5
1008	Tingyi Holding	9.2	0.5	7.5	15
1012	Sun Art Retail Group	12.3	0.4	7.2	13.9
1019	Fosun International	8.2	0.6	26	4.1
1028	Shanghai International Port	4.5	0.8	13.3	9.8
1045	China Cosco Holdings	14	-1.5	26.5	6.4
1073	Shimao Property Holdings	4.5	0.9	19.2	5.9
1111	China Oilfield Services	3.5	0.7	12	11.1
1145	Wuliangye Yibin	3	1	5.9	14.7
1151	Weichai Power	7.6	0.5	10.6	7
1163	Henan Shuanghui Investment	6.3	0.5	2.7	12.6
1165	ZTE	13.3	-0.4	16.4	6.5
1177	Hebei Iron & Steel	21.2	0.2	22.8	4.1
1181	Inner Mongolia Yitai	5	1	6.4	8.7
1187	Bank of Ningbo	2.4	0.5	41.3	4.9
1191	Greentown China Holdings	5.6	0.8	17.3	3.2
1197	Guangzhou R&F	4.8	0.9	15.8	4.8
1218	Qingdao Haier	12.6	0.5	8	5.4

1227	Want Want China	3.4	0.6	3.5	19.8
1246	Bank of Nanjing	2.2	0.5	44.7	4.4
1262	Sino-Ocean Land Holdings	4.6	0.7	20.2	3.5
1288	Huadian Power International	9.4	0.2	26.4	4.8
1312	China Merchants Property	3.3	0.5	17.5	6.1
1322	Jiangsu Yanghe Brewery	2	0.6	2.9	12.4
1333	China Longyuan Power	2.7	0.4	17.3	7.2
1337	GF Securities	1	0.3	12.2	13.9
1342	Wuhan Iron & Steel	16.1	0.2	15.5	4.5
1347	Agile Property Holdings	4.8	0.8	14.7	4
1371	Soho China	2.5	1.7	12.6	4
1385	Inner Mongolia Rare-Earth	1.8	0.6	2.2	11.8
1388	Shenzhen Overseas	3	0.6	11.7	6.7
1418	Shanghai Construction	14.5	0.3	13.3	2.7
1431	BYD	7.4	0	11	8.3
1436	Angang Steel	12.3	-0.7	16.2	4.1
1443	Shanxi Taigang Stainless	15.3	0.3	10.5	3.1
1446	Yang Quan Coal Industry	8	0.4	4.4	5.4
1447	Hengan International Group	2.4	0.5	3.8	11.9
1455	Shanghai Pharmaceuticals	10.8	0.3	8.1	5.6
1485	China National Chemical	6.9	0.4	7.2	6.7
1506	New Hope Liuhe	11.4	0.4	3.5	3.6
1513	BBMG	5.1	0.5	13.3	4.4
1519	Huatai Securities	1.3	0.3	13.6	9.5
1528	Xiamen C&D	12.8	0.4	8.2	2.3
1529	China International Marine	8.6	0.3	10.3	4.9
1534	China Merchants Securities	1.1	0.3	10.9	9.7
1544	Hunan Valin Steel	11.7	0	13.6	1
1553	China Communications Services	9.8	0.4	7.2	4.7
1568	Dongfang Electric	6	0.3	12.6	4.7
1592	Inner Mongolia Yili	6	0.3	3.2	8.9
1610	Maanshan Iron & Steel	11.8	-0.6	12.2	2.3
1613	TCL Corp	11	0.1	12.8	3.4
1613	Shandong Gold-Mining	6.3	0.3	2	7.6
1631	Netease	1.3	0.6	3.1	6.8
1631	Jizhong Energy Resources	6	0.5	6	4.9
1648	Minmetals Development	25.7	0.1	7.6	2.6
1680	Gemdale	3.8	0.5	14.3	4.3
1734	China Gezhouba	8.3	0.2	12.3	2.9
1745	Great Wall Technology	15.2	0	8.5	0.2
1773	Luzhou Laojiao	1.3	0.5	1.9	6.6
1833	Hainan Airlines	4.4	0.3	14.9	4.5
1833	China Hongqiao Group	3.9	0.9	7.1	3

1847	Zhongjin Gold	5.3	0.3	3	6.9
1864	Shanxi Lu'an Environmental	3.1	0.4	6.4	6.7
1866	Qinghai Salt Lake	15.2	0.4	4.7	6.8
1906	Inner Mongolia Baotou Steel	6.8	0.1	7.9	6.4
1916	Shanghai Material Trading	12.8	0	1.5	0.3
1944	Tongling Nonferrous Metals	12.2	0.1	5.9	3.9
1948	Hikvision	0.8	0.2	1.3	12.3