BUSINESS INTELLIGENCE

Business intelligence (BI) refers to the processes, technologies, and tools used by organizations to analyze data and gain insights that inform decision-making. This can include data mining, reporting, dashboards, and data visualization tools, among others. The goal of BI is to provide decision makers with the information they need to make informed decisions and improve overall performance

INFORMATION GATHERING

1. Surveys and Questionnaires: collecting data through surveys and questionnaires to gather information directly from customers and employees.
2. Social Media Monitoring: analyzing data from social media platforms to understand customer sentiment and preferences.
3. Web Analytics: the measurement and analysis of website traffic to gain insights into user behavior and improve website performance.

Data Mining: using statistical and machine learning algorithms to identify patterns and relationships in large data sets.

Data Warehousing: the process of centralizing and storing data from various sources into a single, optimized repository.

OLAP (Online Analytical Processing): a technology that enables multidimensional analysis of data, allowing users to pivot and slice data in real-time.

Dashboards and Visualization: interactive tools that help users quickly understand trends and patterns in data.

Report Generation: the process of creating reports and documents that provide information on key business metrics and KPIs.

BUSINESS INTELLIGENCE REQUIREMENTS

1. Data: Access to high-quality, relevant, and up-to-date data from various internal and external sources.
2. Technology: The right technology stack, including software tools and platforms, to collect, store, and analyze data.
3. Infrastructure: A robust and secure infrastructure to support data storage and processing, with scalability to accommodate growth.
4. Skills: A team with the skills and expertise to manage the BI process, including data engineers, data scientists, and business analysts.
5. Processes: A clear and defined process for data gathering, analysis, and decision-making, with regular audits and performance evaluations.
6. Communication: Effective communication and collaboration between IT, business, and data stakeholders to ensure alignment and buy-in.
7. Governance: A governance framework to ensure data quality, privacy, and security, and ensure compliance with relevant regulations and standards.
8. Strategy: A clear and comprehensive BI strategy that aligns with the overall business strategy and objectives.

BUSINESS INTELLIGENCE STARATAGIES

Data-Driven Decision Making: Using data and insights to inform and drive decision making, rather than relying on intuition or gut feeling.

1. Customer Analytics: Gathering and analyzing customer data to gain insights into customer behavior, preferences, and buying patterns.

2. Predictive Analytics: Using statistical models and machine learning algorithms to predict future outcomes based on historical data.
3. Real-Time Analytics: Processing and analyzing data in real-time to make quick and informed decisions.
4. Big Data Analytics: Handling and analyzing large and complex data sets to uncover hidden patterns and trends.
5. Mobile BI: Enabling mobile access to business intelligence and analytics, allowing users to access insights and make decisions on the go.
6. Social Media Analytics: Analyzing data from social media platforms to understand customer sentiment and preferences, and improve customer engagement.
7. Collaborative BI: Encouraging collaboration and teamwork between business and IT teams to drive data-driven decision making.
8. Cloud BI: Utilizing cloud-based solutions for data storage, processing, and analysis, providing scalability, flexibility, and cost savings.
9. Self-Service BI: Empowering end-users to access and analyze data without relying on IT or data experts, increasing agility and speed.

STRATEGIC APPROACH TO BI

1. Define business objectives: Clearly articulate the goals and objectives of the organization and align BI initiatives with these goals.
2. Assess data needs: Identify the data required to support decision making and determine how this data can be obtained.
3. Choose the right technology: Select the right technology stack, including software tools and platforms, to support data collection, analysis, and visualization.
4. Develop a data strategy: Define a data strategy that outlines how data will be collected, stored, and analyzed to support decision making.
5. Implement the right processes: Develop and implement processes for data gathering, analysis, and decision-making that are efficient and effective.
6. Encourage collaboration: Foster collaboration between IT, business, and data stakeholders to ensure alignment and buy-in.

7. Monitor performance: Regularly monitor and evaluate the performance of the BI implementation and make adjustments as needed.
8. Continuously improve: Continuously refine the BI strategy and processes to ensure they remain relevant and effective.

BI AND BA

BI	Business Analytics
Business intelligence analyzes present and past data to understand what happened.	Business analytics evaluates past data to predict what will happen in the future. It's related to long-term operations.
You need BI to run daily operations by tracking present performance.	Business analytics helps improve operations and productivity organization-wide, focusing on future strategy.
Business intelligence is descriptive with tables and reports of business performance data.	Business analytics is complex, incorporating diagnostic, predictive and prescriptive analytics.
BI supports decision-making by answering specific queries and viewing key metrics.	Business analytics is open-ended. You can dive deeper as you go, posing iteratively detailed questions to get the desired insight. Business analytics usually has an eye toward improvement and preparation for change.
Its applications include performance management through visual insight and reporting.	Its applications include forecasting business performance and market trends through statistical data analysis and modeling.
Power BI, Qlik Sense, MicroStrategy and Dundas BI are BI tools.	Business analytics tools include word processing tools, cloud-based spreadsheets, documents, MS Office tools, etc. Power BI, Qlik Sense, MicroStrategy, and Dundas BI are business analytics and business intelligence solutions.

Data Collection and Binding

Data collection and binding in business intelligence refers to the process of gathering and consolidating data from multiple sources into a single, unified repository for analysis and reporting purposes. The collected data is usually stored in a data warehouse or a data mart and is used to support decision-making, planning and strategy formulation. The data can be in various formats such as structured, semi-structured or unstructured and the binding process involves cleaning, transforming, and integrating the data to ensure consistency and accuracy.

STRUCTURED AND UNSTRUCTURED

Structured data refers to data that is organized in a pre-defined manner and can be easily stored, processed and analyzed by computers. It is typically stored in a database and can be easily searched, sorted, and aggregated. Examples include tables, spreadsheets, and databases.

Unstructured data refers to data that does not have a pre-defined structure and cannot be easily processed by computers. It is often in the form of text, images, audio, and video and is more difficult to search, sort, and analyze. Examples include emails, social media posts, customer reviews, and multimedia files.

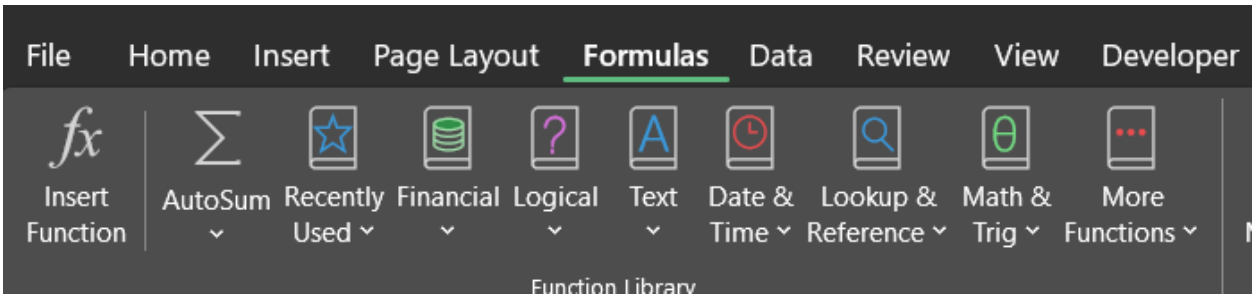
INTRODUCTION TO EXCEL

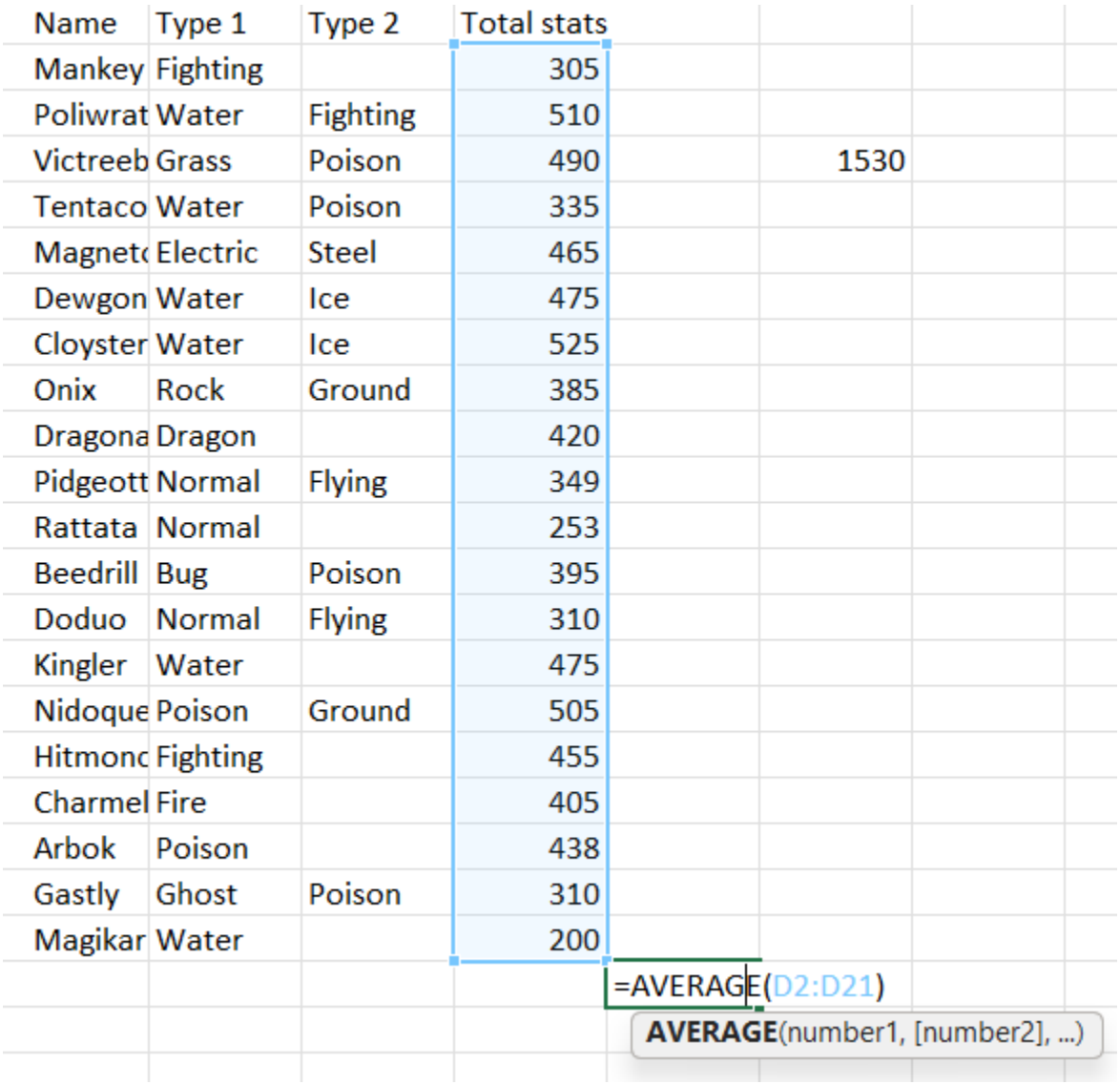
FUNCTIONS

Microsoft Excel has over **400 built-in functions**. These functions cover a wide range of categories such as mathematical and statistical, financial, date and time, text, and database functions. The functions can be used to perform complex calculations, analyze data, and automate tasks, among other things.

Functions actually help to understand the complex and large datasets in quicker way. its a logical approach to understand and analyze the data.

in **Formulas** menu you can find the formulas/functions Ribbon (see. Below Image)





Name	Type 1	Type 2	Total stats
Mankey	Fighting		305
Poliwrat	Water	Fighting	510
Victreeb	Grass	Poison	490
Tentaco	Water	Poison	335
Magnet	Electric	Steel	465
Dewgon	Water	Ice	475
Cloyster	Water	Ice	525
Onix	Rock	Ground	385
Dragona	Dragon		420
Pidgeott	Normal	Flying	349
Rattata	Normal		253
Beedrill	Bug	Poison	395
Doduo	Normal	Flying	310
Kingler	Water		475
Nidoque	Poison	Ground	505
Hitmonc	Fighting		455
Charmel	Fire		405
Arbok	Poison		438
Gastly	Ghost	Poison	310
Magikar	Water		200

👉 To apply Function to the data we need to start with **“=” (Equal to symbol) then Function name(in brackets apply cell names according to syntax). Refer the above picture.**

List Functions with Syntax

- `AVERAGE(number1, [number2, ...])`: Returns the average of a set of numbers.

- `COUNT(value1, [value2, ...])`: Returns the number of cells that contain numbers in a range.
- `MAX(number1, [number2, ...])`: Returns the largest value in a set of values.
- `MIN(number1, [number2, ...])`: Returns the smallest value in a set of values.
- `SUM(number1, [number2, ...])`: Returns the sum of a set of numbers.
- `MEDIAN(number1, [number2, ...])`: Returns the median (middle value) of a set of numbers.
- `MODE(number1, [number2, ...])`: Returns the most frequently occurring value in a set of numbers.
- `ABS(number)`: Returns the absolute value of a number.
- `SQRT(number)`: Returns the square root of a number.
- `POWER(number, power)`: Returns the result of a number raised to a power.
- `EXP(number)`: Returns e raised to the power of a number.
- `LN(number)`: Returns the natural logarithm of a number.
- `LOG10(number)`: Returns the base-10 logarithm of a number.
- `ROUND(number, num_digits)`: Rounds a number to a specified number of digits.
- `CEILING(number, significance)`: Rounds a number up to the nearest multiple of significance.
- `FLOOR(number, significance)`: Rounds a number down to the nearest multiple of significance.
- `INT(number)`: Rounds a number down to the nearest integer.
- `MOD(number, divisor)`: Returns the remainder after a number is divided by a divisor.
- `RAND()`: Returns a random number between 0 and 1.
- `RANDBETWEEN(bottom, top)`: Returns a random integer between bottom and top, inclusive.
- `RANK(number, ref, [order])`: Returns the rank of a number in a set of values.
- `PERCENTRANK(array, x, [significance])`: Returns the rank of a value in a set of values as a percentage of the set.

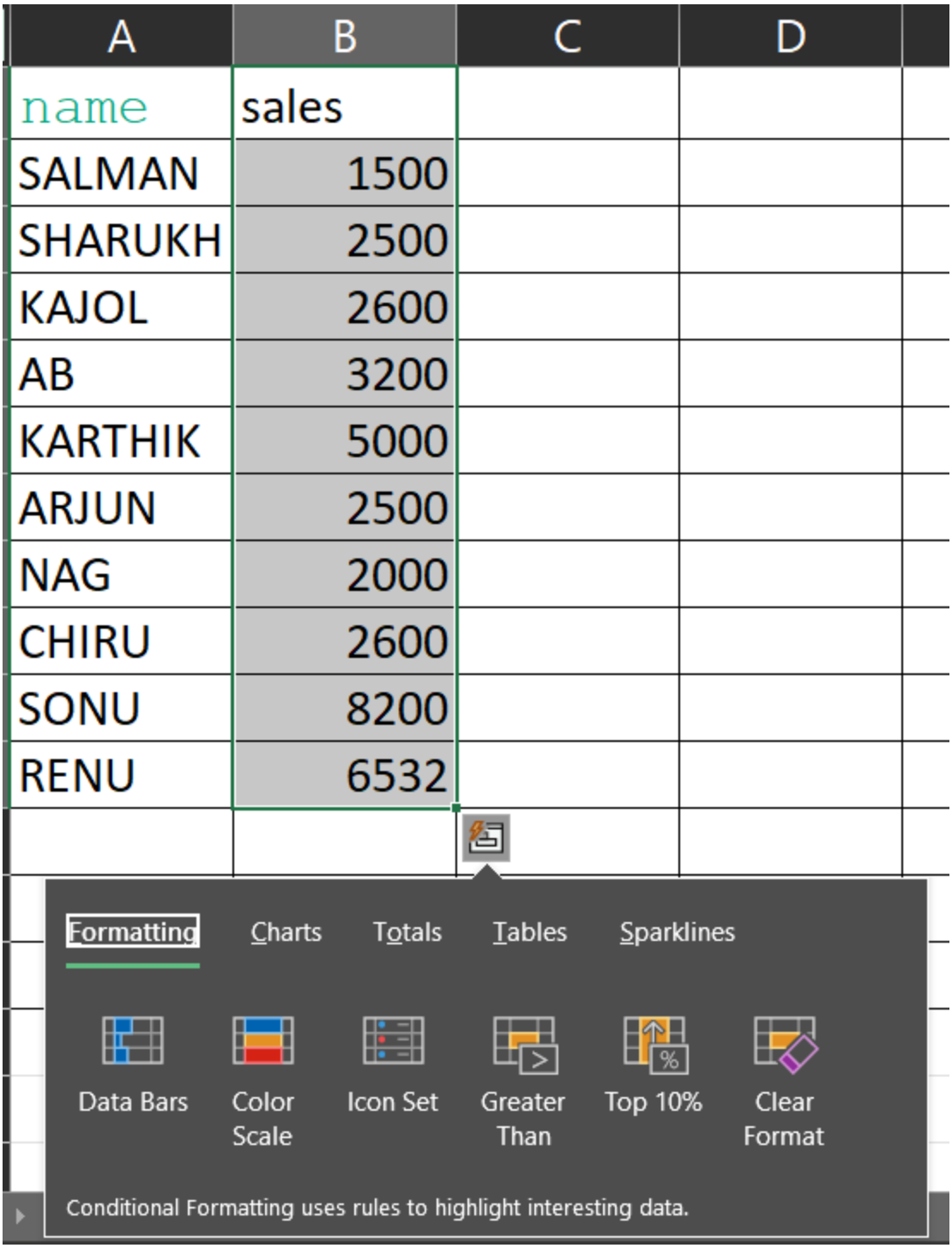
- `PERCENTILE(array, k)`: Returns the k-th percentile of a set of values.
- `STDEV(number1, [number2, ...])`: Returns the standard deviation of a set of numbers.
- `STDEVP(number1, [number2, ...])`: Returns the standard deviation of a population of numbers.
- `VAR(number1, [number2, ...])`: Returns the variance of a set of numbers.
- `VARP(number1, [number2, ...])`: Returns the variance of a population of numbers.
- `COVAR(array1, array2)`: Returns the covariance, a measure of the relationship between two sets of numbers.
- `CORREL(array1, array2)`: Returns the correlation coefficient, a measure of the relationship between two sets of numbers.
- `FORECAST(x, known's, known)

Charts

Charts are used to represent the data in excel. Excel supports various chart types such as **column charts**, **bar charts** , **line charts**, **pie charts**, **scatter charts** and others. You can customize the appearance of the charts and graphs, including colors, data labels, axis labels, and legends.

Quick Analysis (Ctrl +Q)

- **Quick Analysis:** The Quick Analysis tool provides a quick way to apply basic formatting, charts, and sparklines to data in a worksheet. Simply select the data you want to analyze and then click on the Quick Analysis button to access its features.
-



- **Pivot Tables**: A pivot table is a powerful tool for summarizing and aggregating data. You can quickly create a pivot table from a range of data and then manipulate it to see different summaries and groupings.

[LINK]

- **Conditional Formatting:** You can use conditional formatting to apply visual styles to data based on specific conditions, such as highlighting cells that contain values above or below a certain threshold.
- **Data Sorting:** You can sort data in Excel by column to quickly reorganize it in ascending or descending order.
- **Data Filtering:** You can use the filter function to display only the data you want to see, based on specific criteria.

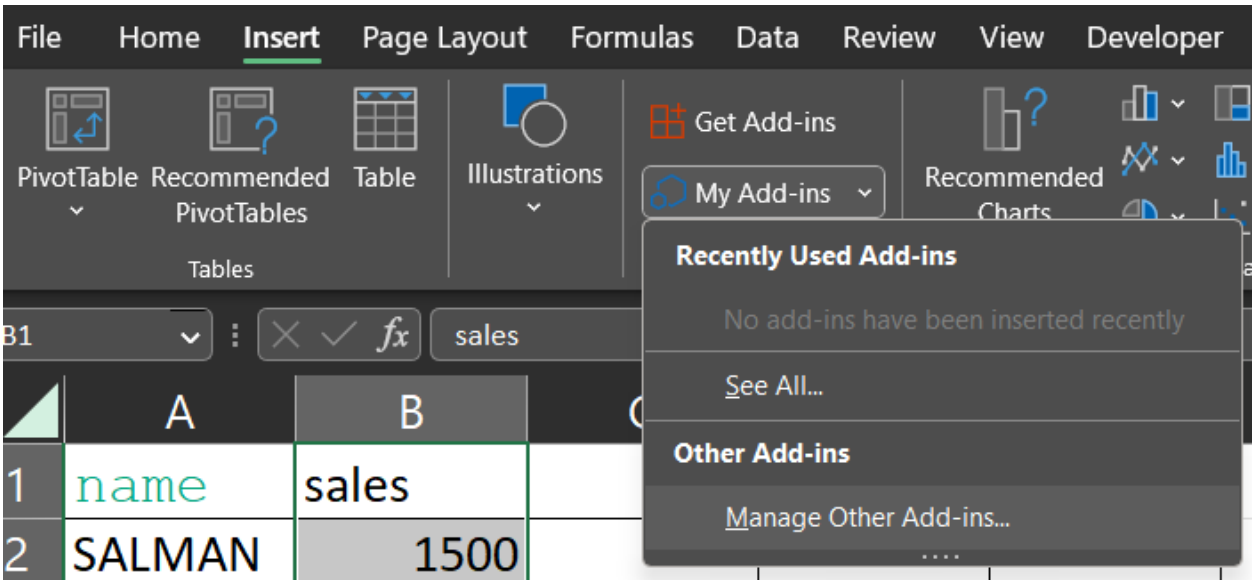
DATA ANALYSIS TOOL PACK

The Data Analysis Tool pack is an add-in for Microsoft Excel that provides a set of tools for more advanced data analysis. The tool pack includes:

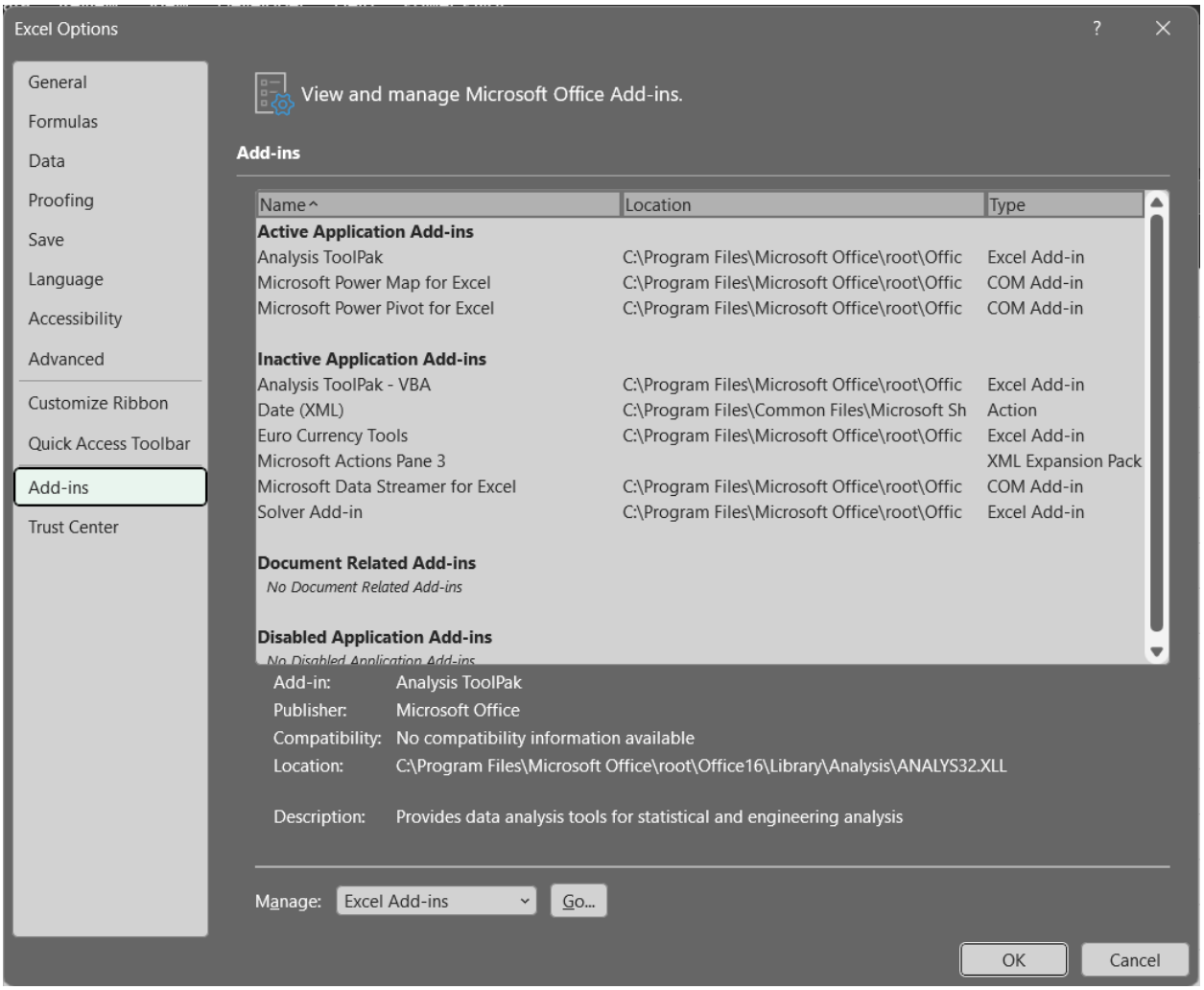
> 👉 The Data Analysis Tool pack is a valuable tool for data analysis and helps you to quickly and easily perform more advanced statistical analysis on your data

1. Descriptive Statistics: Provides basic statistical measures such as mean, median, mode, standard deviation, and more.
2. Histogram: Generates a histogram, which is a bar graph that shows the distribution of data.
3. Sampling: Draws a random sample from a population of data.
4. Regression: Analyzes the relationship between two sets of data and fits a line of best fit through the data points.
5. ANOVA (Analysis of Variance): Tests for differences between the means of multiple groups of data.
6. Exponential Smoothing: Forecasts future values based on a weighted average of past data.
7. Moving Average: Calculates the average of a set of values over a specified number of periods.

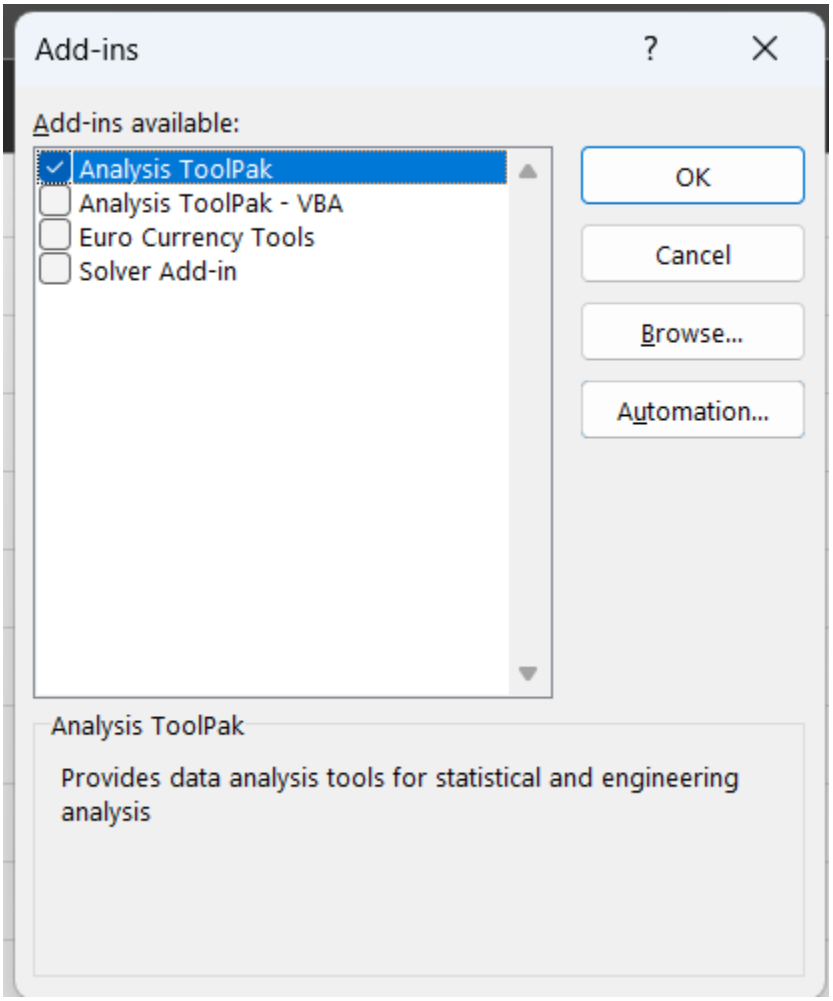
How to Activate Data Tool Pack



`Click on → Mange Other Add-ins..`

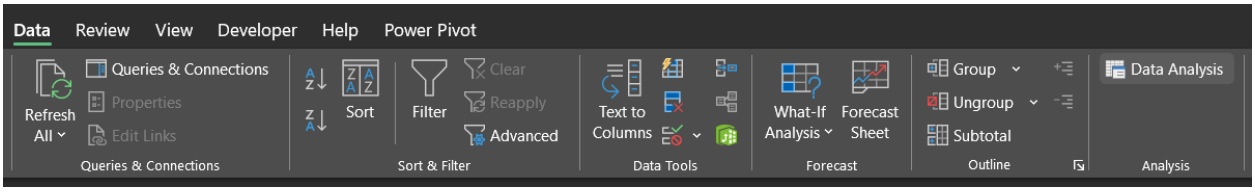


`Mange → Excel Add-ins(Default) → Click go`

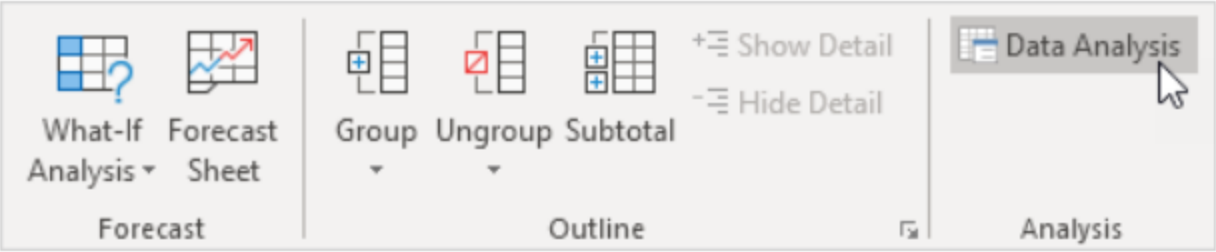


`Check the tick mark to Analysis Toolpak click ok`

👉 To access the tool pack, go to the Data tab in the ribbon, click on the Data Analysis button, and select the analysis tools

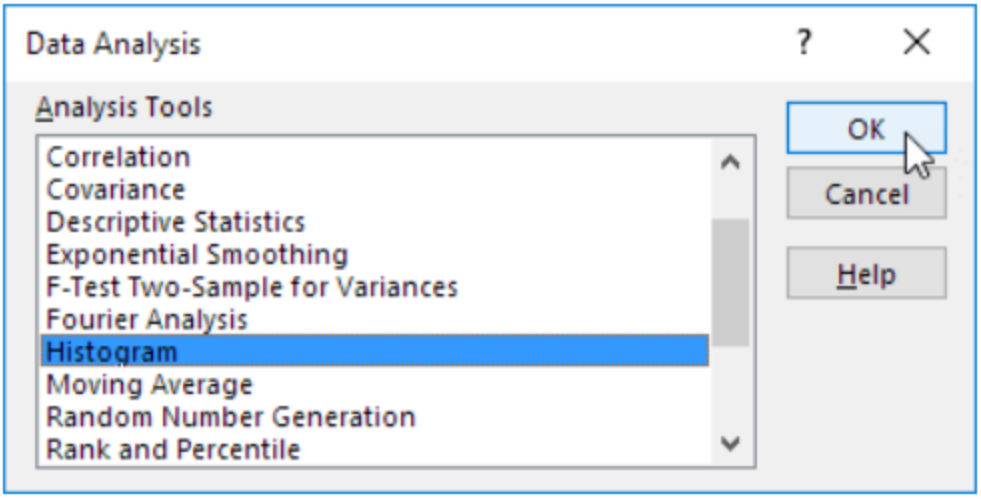


To access

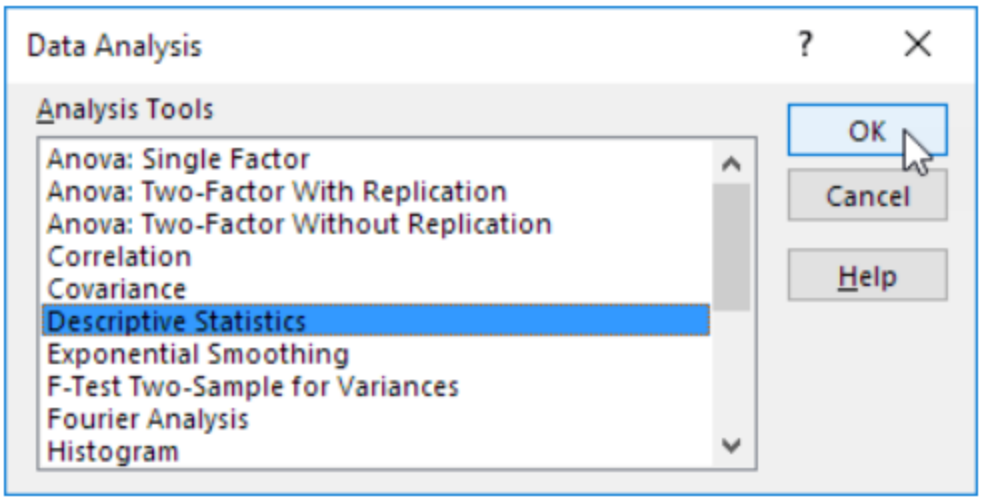


The following dialog box below appears.

5. For example, select Histogram and click OK to create a Histogram in Excel.



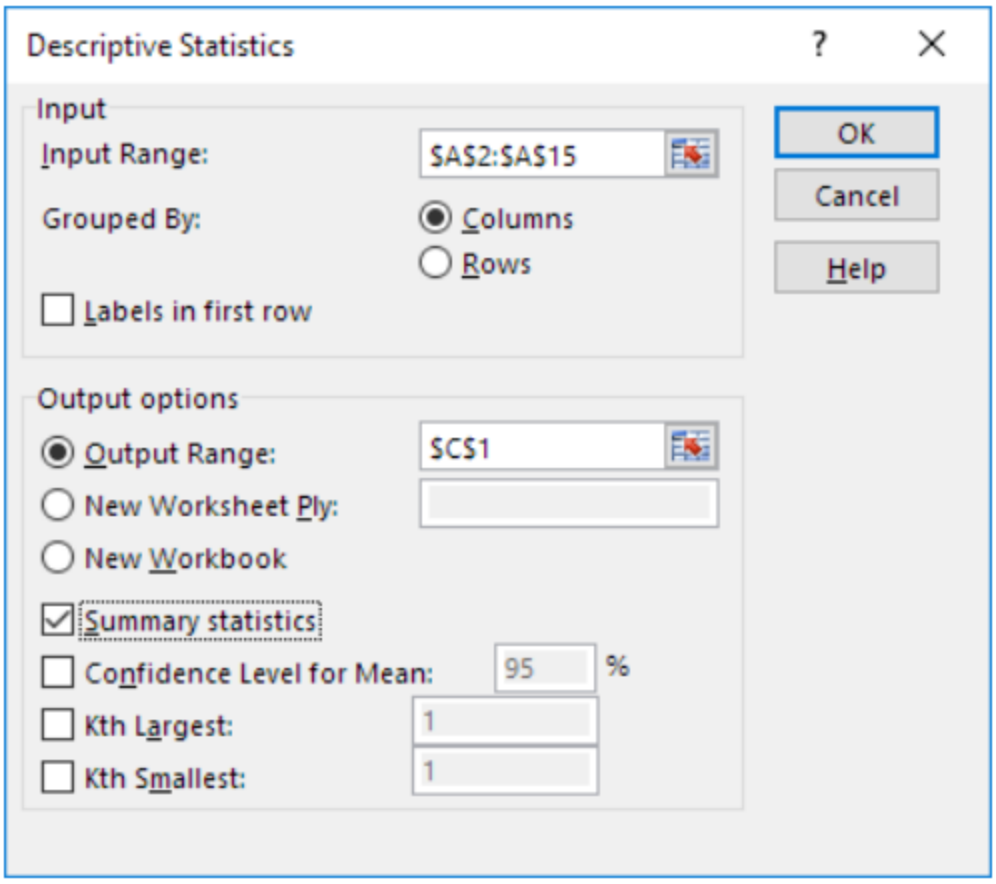
To work on Descriptive Summaries in Data analysis



3. Select the range A2:A15 as the Input Range.

4. Select cell C1 as the Output Range.

5. Make sure Summary statistics is checked.



OUTPUT

Result:

	A	B	C	D	E
1	Scores		*Column1*		
2	82				
3	93		Mean	81.21428571	
4	91		Standard Error	4.045318243	
5	69		Median	85	
6	96		Mode	93	
7	61		Standard Deviation	15.13619489	
8	88		Sample Variance	229.1043956	
9	58		Kurtosis	-1.426053506	
10	59		Skewness	-0.402108004	
11	100		Range	42	
12	93		Minimum	58	
13	71		Maximum	100	
14	78		Sum	1137	
15	98		Count	14	
16					

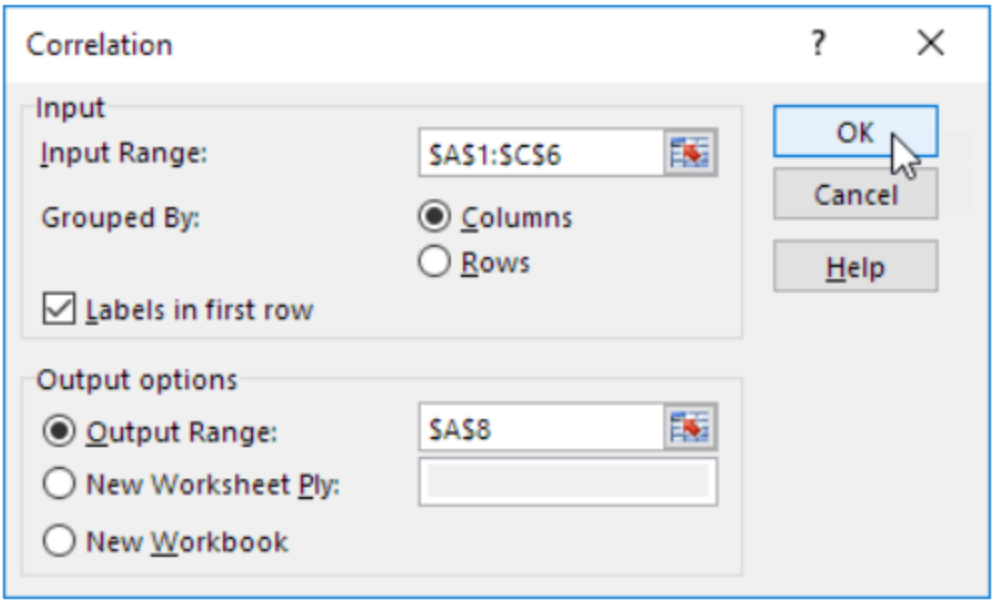
CORRELATION

The correlation coefficient (a value between -1 and +1) tells you how strongly two variables are related to each other. We can use the CORREL function or the Analysis Toolpak add-in in Excel to find the correlation coefficient between two variables.

- A correlation coefficient of +1 indicates a perfect positive correlation. As variable X increases, variable Y increases. As variable X decreases, variable Y decreases.
- A correlation coefficient of -1 indicates a perfect negative correlation. As variable X increases, variable Z decreases. As variable X decreases, variable Z increases.

To find out the correlation between two variables

Data Menu → Data Analysis → select correlation



Fill the required data fields with data point in excel sheet

Output of the correlation (matrix format)

	A	B	C
A	1		
B	0.191516	1	
C	0.909268	0.108893	1

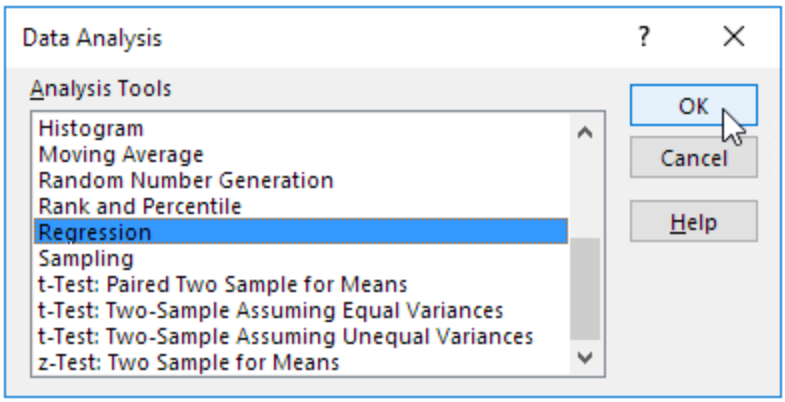
REGRESSION

Regression is a statistical method used to analyze the relationship between two variables. It aims to find the line of best fit that summarizes the relationship between the variables. The line of best fit represents the average change in the dependent variable (also known as the "outcome" or "response" variable) for every unit change in the independent variable (also known as the "predictor" or "explanatory" variable).

For example, if you wanted to determine the relationship between the number of hours studied and the score on a test, you could use regression to find the line of best fit that summarizes the relationship between hours studied and test score. The line of best fit

could show that, on average, a one-hour increase in study time results in a two-point increase in test score.

Regression can be used to make predictions about the outcome variable based on the predictor variable. It can also be used to identify the strength and direction of the relationship between the variables.



3. Select the Y Range (A1:A8). This is the predictor variable (also called dependent variable).

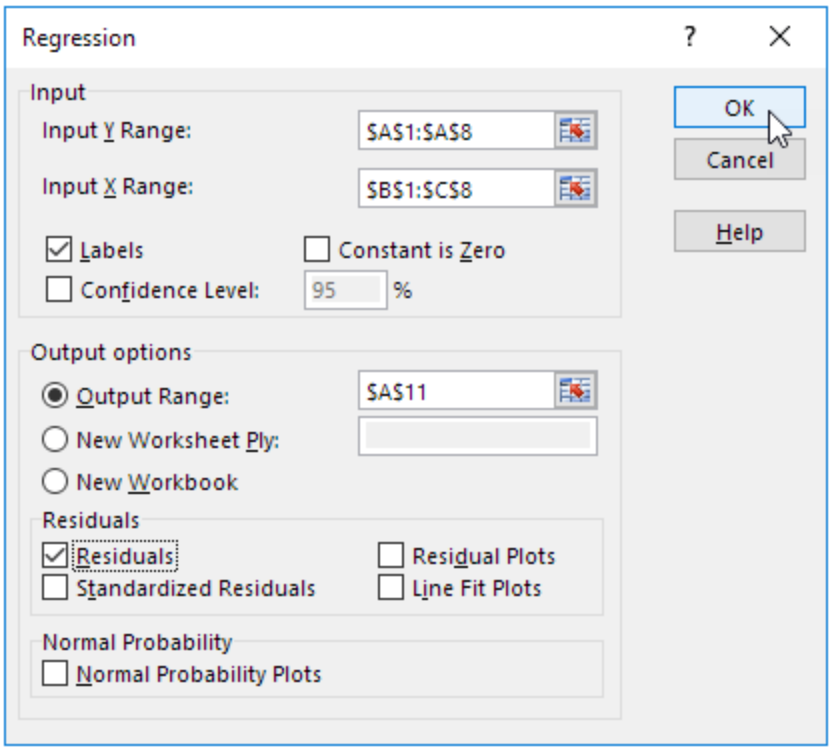
4. Select the X Range(B1:C8). These are the explanatory variables (also called independent variables). These columns must be adjacent to each other.

5. Check Labels.

6. Click in the Output Range box and select cell A11.

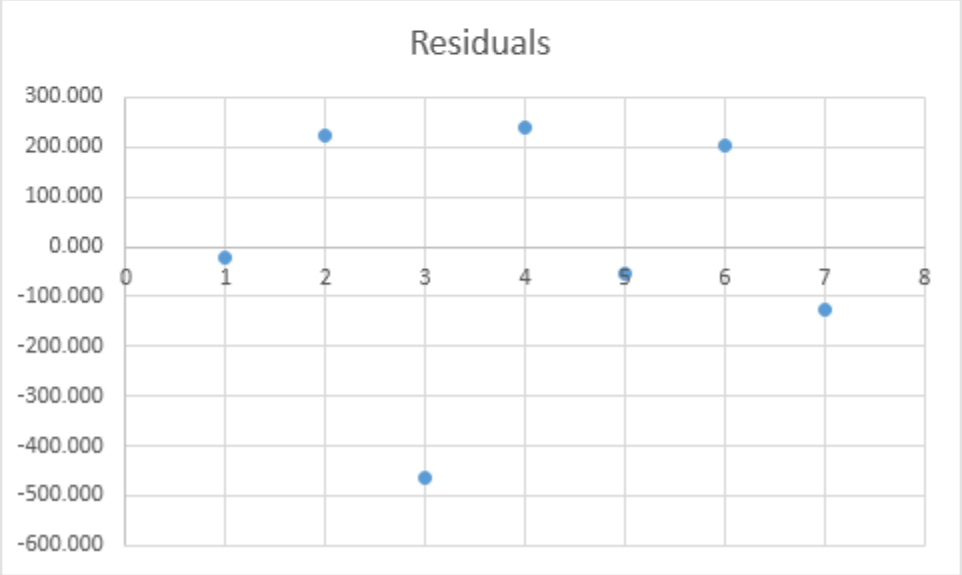
7. Check Residuals.

8. Click OK.



33	RESIDUAL OUTPUT		
34			
35	*Observation*	*Predicted Quantity Sold*	*Residuals*
36	1	8523.009	-23.009
37	2	4476.048	223.952
38	3	6265.938	-465.938
39	4	7160.883	239.117
40	5	6252.733	-52.733
41	6	7095.058	204.942
42	7	5726.330	-126.330
43			

You can also create a scatter plot of these residuals.



Data Analytics