
40
DAYS

Data Analysis Task

Data analytics is an essential part of any business workflow nowadays. So here some tools to process big data :

- MongoDB
- R Programming Environment
- Neo4j
- Apache Hadoop

What is the difference between a data analyst and a data scientist?

Data analysts work on the simpler parts whereas Data scientists work on more complex parts.

Data Analysts

- Focus on statistics and mathematics
- Required programming langues :R , sql .HTML ,Javascript
- Must be able to able analyse and have cognizance of spreadsheet tools like : Excel and Power BI
- Should have knowledge of data visualization : tableau
- Qualification needed : Should be familiar with data wharehousing ,Business intelligence and analysis in SQL.
- Qualification needed : Should know about data storing , retrieving skills
- Responsibilities :

1-Applying statistical analysis

2- Designing and creating data reports to help stakeholders make better decisions

3-Analyzing and mining business data to discover patterns and identify correlations from different data points

Data Analysis Task

Data Scientist

- Focus on statistics and mathematics
- Required programming langues :R , sql .HTML ,Matlab ,SAS
- Business Acumen
- Should have knowledge of data visualization : tableau
- Qualification needed : Should be familiar with all database system like MYSQL.
- Qualification needed : Should know about collecting data and using them on data sets
- Should have knowledge about machin learning algorithm
- Responsibilities :

1-Cleansing and processing of data
2- Developing machine learning models and new analytical methods
3-Data visualization and storytelling

What is data visualization?

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.
In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

Data Analysis Task

What is data visualization?

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.
In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.
Common general types of data visualization:

- Charts
- Tables
- Graphs
- Maps
- Infographics
- Dashboards

Some tools of data visualization

1-Tableau

2-Adaptive Discovery

3-JupyteR

4-Dundas BI

5-Google Charts

Data Analysis Task

Why data visualization is important for business decisions:

1- Provides Multiple Perspectives for Analysis

Businesses often rely on spreadsheets or other static representations like charts or graphs of their data. The pitfall here is that a single representation might mean a higher possibility of misinformation.

2- Assists Faster Action on Emerging Trends

3-Aiding a Unified Business Vision

4- Improved response times

Data visualization puts the data into the users' hands allowing them to more quickly identify issues and improve response times.

5- Greater Simplicity

Using visualizations allows users to get the big picture and see the details at the same time.

6- Easier visualization of patterns

It allows decision-makers to view data using graphical representations including charts, fever charts, and heat maps.

7- Enhanced Collaboration

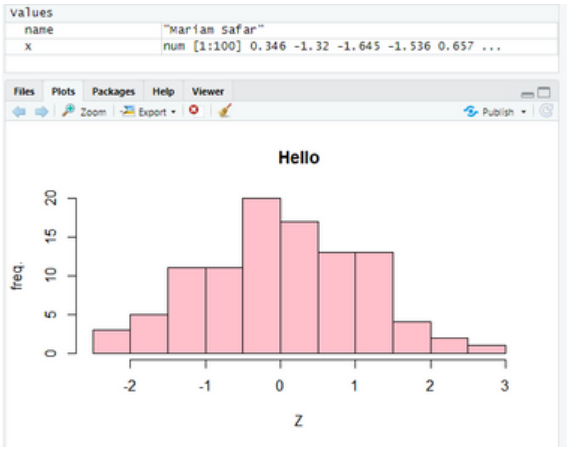
Collaboration gives teams access to the strengths and skills of everyone involved by combining the experiences of the entire group. These skills can be used to solve problems faster and improve innovation. Using advanced visualizations makes it easier for teams to collaborate.

Data Analysis Task

R language :

Simple attempts

```
> x <- c("hello world")
> (x)
[1] "hello world"
```




```
mydata=read.csv("new.csv",TRUE,",")
mydata
```

```
mydata=read.csv("new.csv",TRUE,",")
mydata
  Name.Gender.Color
    Maha;female;red
     Ali;male;green
    Ahmad;male;blue
Fatimah;female;pink
   Saeed;male;black
Maria;female;yellow
```