Summary – Day 1

- Tableau is a leading data visualization tool used for data analysis and business intelligence to get the insights out of the data.

- Below are the products of Tableau.
  - **Desktop**: Used to create the charts & dashboards.
  - **Server**: Used to publish the dashboards, create the users, schedule dashboards and datasets, give permission to the users, create user groups.
  - **Online**: Used to perform the same tasks as above. The difference is the online is cloud based platform.
  - **Prep**: Used to perform the ETL (Extract, Transform, Load) operation on the source data so that all the incorrect/invalid values can be transformed.
  - **Reader**: Used to view the others' workbooks. Cannot modify those.
  - **Public**: It is an open-source website on which many tableau enthusiasts are active to share and collaborate their personal tableau work with others.

- Data visualisation is a concept used to visualise the data in pictorial format by creating the charts, dashboards, rather than looking at the same data in traditional rows/columns format. Used to carry out the data analysis effectively.

- Tableau Desktop is used by data analysts to create the dashboards. Tableau Developer, BI Developer, Data Analyst, Business Analyst are some job roles associated with using tableau desktop.

- It is a licensed software which comes with below three variations.
  - **Professional**: Free for 14 days on trail basis. Post that we need to activate it using product key.
  - **Professional -Student**: Free for 1 year only for students who have valid student ID card. Post 1 year we need to redo the same activity.
  - **Public**: Free forever. No licensing required. However, there are some features/operations missing here w.r.t. professional edition.

- Tableau Professional (14 days free)
  https://www.tableau.com/products/desktop/download

- Tableau Student Version (1 year free)
  https://www.tableau.com/academic/students

- Tableau Public (Free forever)
  https://www.tableau.com/products/public/download

- Below are the different file extensions available in tableau.
  - **Tableau Workbook (.twb)**
    - Need to send the source data and tableau file to the sender in order to open it. If only .twb file is sent the then other user will not be able to open it. In fact, user will receive an error.
  - **Tableau Packaged Workbook (.twbx)**
    - Single zip file containing a workbook along with any supporting local file data and background images. Here first person just needs to send the workbook in .twbx file format. No need to send the data additionally.
  - **Extract (.hyper)**
    - Local copy of a subset or entire data set that can be used to share data with others, when need to work offline, and improve performance of tableau dashboards.
  - **Data Source (.tds)**
    - Data source files are shortcuts for quickly connecting to the original data that you use often. Data source files do not contain the actual data but rather the information necessary to connect to the actual data.
  - **Packaged Data Source (.tdsx)**
    - A packaged data source is a zip file that contains the data source file (.tds) described above as well as any local file data such as extract files (.hyper), text files, Excel files, Access files, and local cube files.
  - **Bookmarks (.tbm)**
    - Contains a single worksheet and is an easy way to quickly share your work. E.g. If some complex chart needs to be created then instead of recreating it again and again in different workbooks, we can create it once and then by saving the file in .tbm format we can open it in a new workbook requiring that chart.

- Below are basic data types available in tableau.
  - **String**: Deals with plain text
  - **Numerical**: Deals with numbers
  - **Date**: Date data in day, month and year format
  - **Date-time**: Date data in day, month, year, hour, minute and second format
  - **Boolean**: True or False values based on condition

- **Geographical**: Geo-spatial data i.e. country, state etc.

- Dimensions are discrete, qualitative data which deals with finite values e.g. customer name, country, 01-01-2023 etc.

- Measures are continuous, quantitative data which deals with infinite values e.g. sales, profit etc.

- Aggregation is the concept of combining the values of either dimension or measure and applying aggregate functions on top of it. e.g. SUM(Sales), AVG(Profit), COUNT(Cust name)
- There are 2 types of connections Live and Extract.
  - **Live**:
    - Real-time data flowing form original source in tableau.
    - Live data processes queries in source database.
    - By default, connection is live in tableau.
    - With live connection data in the worksheet gets updated every time when it is opened.
  - **Extract**:
    - Tableau Data Extract(.tde/also referred to as hyper file) is a compressed snapshot of live data that is stored locally and loaded into memory.
    - To ease off our connection and minimize the load due to complex data, we can apply filters to our data, save it as a .hyper file and later use them for our extract connection.
    - This tends to be much faster than Live Connections and hence provides an edge when used for complex visualizations. Manually need to click on Refresh option.