Data Analysis Fundamentals
Lecture 7: Power Query for Data Analysis

Vikas Kumar
Indian Institute of Technology Bombay
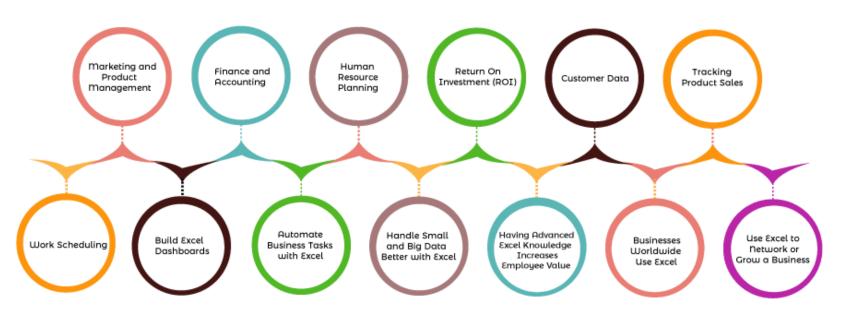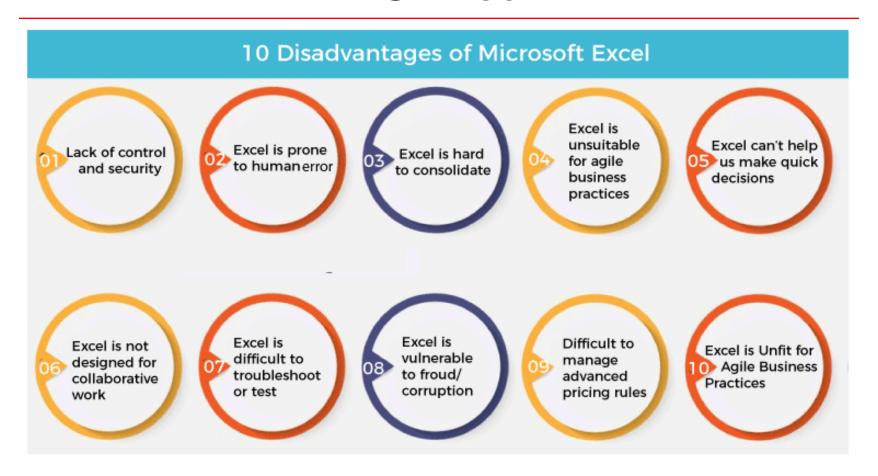
12th September 2023

Outline

- ➢ Challenges of working with Excel

- ➢ Graduating from Excel to PowerQuery

- ➢ Overview of PowerQuery as a matured tool for Data Analysts

- ➢ Applying all concepts learned till date in a mini in-class project

MS Excel

Why MS Excel is Important in Business

Marketing and Product Management

Finance and Accounting

Human Resource Planning

Return On Investment (ROI)

Customer Data

Tracking Product Sales

Work Scheduling

Build Excel Dashboards

Automate Business Tasks with Excel

Handle Small and Big Data Better with Excel

Having Advanced Excel Knowledge Increases Employee Value

Businesses Worldwide Use Excel

Use Excel to Network or Grow a Business

MS Excel

10 Disadvantages of Microsoft Excel

01 Lack of control and security
02 Excel is prone to human error
03 Excel is hard to consolidate
04 Excel is unsuitable for agile business practices
05 Excel can't help us make quick decisions
06 Excel is not designed for collaborative work
07 Excel is difficult to troubleshoot or test
08 Excel is vulnerable to froud/ corruption
09 Difficult to manage advanced pricing rules
10 Excel is Unfit for Agile Business Practices

Power Query

➢ Power Query is a self-service ETL (Extract, Transform, Load) tool which runs as an Excel add-in.

➢ Feature of Microsoft Excel that was introduced as an add-in to Excel 2010 and 2013 and is now a native feature for Excel 2016 and 365.

➢ Enables to import millions of rows of data from multiple data sources into a single Excel workbook

➢ Create relationships between heterogeneous data

➢ Create calculated columns and measures using formulas

➢ Build PivotTables and PivotCharts for data analysis

Power Query

➢ Import and manipulate hundreds of millions of rows of data where Excel has a hard constraint of just over a million rows.

➢ Import data from multiple sources into one single source workbook without having to create multiple source sheets that suffer from version control and transferability issues.

➢ Manipulate the imported data, analyze it, and draw conclusions without slowing down your computer.

➢ Visualize the data with PivotCharts and Power BI.