DATA ANALYST ROADMAP 2.0 (4 Months): For 2023 - 24

*Note – The time period is applicable if you study for 6-8hrs/day

Data analytics is a rapidly growing field that focuses on extracting insights and value from data to drive informed decision-making. As businesses and organizations collect vast amounts of data, there is a growing demand for professionals who can effectively analyse and interpret that data to gain meaningful insights.

As a data analyst, there are several key skills and areas of knowledge you should focus on to excel in your career. Here's a roadmap to help you navigate your journey as a data analyst in just 3 months:

TOPIC 2 - STATISTICS FOR DATA ANALYSTICS

Statistics plays a crucial role in the field of data analysis. As a data analyst, understanding statistical concepts and techniques allows you to make sense of data, draw meaningful insights, and make informed decisions

- Measures of central tendency: avg/mean, median, mode
- Aggregated Functions: sum, avg, min, max, pct, etc.
- Measures of dispersion: range, variance, standard deviation
- Percentiles and quartiles
- Time series data and components
- Smoothing techniques: moving averages, exponential smoothing
- Gaussian & Normal Distributions

Resources/ Study Links

Statistics -
https://www.youtube.com/watch?v=_D7KHkH15ds
https://youtu.be/LZzq1zSL1bs (First Two and half Hours)

TOPIC 2 - EXCEL

MODULE 1: EXCEL

Excel is a powerful tool for data analysts and is widely used for data manipulation, analysis, and visualization. Here are some key areas to focus on to excel in Excel for data analytics:

- Basic Excel functions and commands.
- Perform Operations with Functions and Formulas like SUM, MIN, MAX, COUNT, AND, OR, IN, AVG, etc.
- Perform Conditional Operations by using functions like IF, SUMIF, AVERAGEIF, COUNTIF, etc.
- Format and Modify text using functions like RIGHT, LEFT, MID, UPPER, LOWER, CONCATENATE, PROPER, etc.
- Use Advanced calculations like VLOOKUP, HLOOKUP, XLOOKUP, LOOKUP, RANDOM, FILTER, DGIT, GET, INDEX, MATCH, SIZE, ISBLANK, ISERROR, TRIM, STRING FUNCTIONS, etc
- Basic to Advanced Date and Time functions to modify the date in different formats.
- Conditional Formatting, conditional formatting rules, manage rules.
- Charts like Bar chart, Stacked Bar chart, Column Chart, Donut, Pie, Tree Maps, Line chart, Area Chart, Map chart, etc.
- Custom Charts like Waffle Chart, Picture Chart, Lollipop Chart, Waterfall Chart, etc.
- Tricks to minimize the tedious and repetitive tasks in Excel.
- Insert and Format Objects, Shapes like Rectangles, Diamonds, Arrows, etc.
- Custom Data Formats and Layouts like custom numbering, advanced fill options

MODULE 2: MIS (MANAGEMENT INFORMATION SYSTEM)

MIS stands for Management Information System. In the context of data analytics, MIS refers to the use of information systems and technology to gather, process, analyse, and present data for effective decision-making and management.

- Pivot Table creation using different datasets of different domains.
- Exploring the benefits of using pivot tables for summarizing and analysing large datasets.
- Understanding the structure of a pivot table and its components (rows, columns, values, filters).
- Pivot charts and Advanced Pivot charts, lay outing and formatting the charts, styling the charts.
- Insert Slicers, Insert Timelines to change the data dynamically and then change the Chartsdynamically.

- Filed List, Buttons, Advanced custom calculations fields, Filed headers, etc.
- Grouping and Sorting
- Utilizing drill-down functionality to access detailed information behind summarized data.
- Create an Interactive and Dynamic MIS/ Excel Dashboard using different datasets.
- 2 Practice Dashboard from end to end

MODULE 3: EXCEL SKILLS USING CHATGPT

- ChatGPT can help you generate and debug complex formulas that may be needed in your day-to-day work.
- Describe the desired outcome or the logic behind the calculation, and ChatGPT can provide suggestions or help you troubleshoot formula errors.
- Understand and utilize advanced functions and formulas in Excel.

MODULE 4: EXCEL PROJECTS

- Project 1 – End to End Development by Candidate
- Project 2 – End to End Development by Candidate

Domain – Ecommerce, Hotel Management, Sales, Food Supply chain, Purchase orders, etc.

Resources/ Study Links

Excel - https://www.youtube.com/watch?v=-ujVQzTtxSg&list=PLWPirh4EWFpEpO6NjjWLbKSCb-wx3hMql

MIS - https://www.youtube.com/watch?v=4wzWwyj7v2s&list=PLmHi7ol7EPSCcS7mM4xZHIF-tlY8t1s4y (Hindi)
https://www.youtube.com/watch?v=RkQl2wVpQAo (English)

Excel Projects -
https://youtu.be/XeWfLNe3moM
https://youtu.be/4WUzvr9jxGc
https://youtu.be/xHTUP1Dxu-M

Excel Skills with ChatGPT -
https://youtu.be/JYtZ2zsdE_s
https://youtu.be/21bXQDXSxYs
https://youtu.be/qB4_2_YX9uE

TOPIC 3 - DATAWAREHOUSING CONCEPTS

In data analytics, a data warehouse plays a crucial role in storing and organizing large volumes of data from various sources to support analytical processes. A data warehouse is a centralized repository that collects, integrates, and manages structured and semi-structured data from different operational systems within an organization.

- DW Architecture
- Need of DW
- Data Mart, ETL Process
- OLAP, OLTP
- Facts, Attributes and Dimensions
- Types of Dimensions
- Normalization and Denormalization
- Schemas- Star, Snowflake, Galaxy

> ### Resources/ Study Links
>
> **Datawarehouse –**
> https://youtu.be/CHYPF7jxlik
> https://www.youtube.com/watch?v=J326LIUrZM8&list=PL9ooVrP1hQOEDSc5QEbI8WYVV_EbW KJwX

TOPIC 4 - SQL

SQL plays a crucial role in data analytics for data analysts. SQL (Structured Query Language) is a widely used language in data analytics and database management. It provides a standard way to interact with relational databases and is particularly useful for querying, manipulating, and analysing structured data.

MODULE 1: SQL – ORACLE & MS SQL SERVER DB / POSTGRES SQL & MY SQL DB

- Introduction to SQL
- ORACLE Server Architecture
- ORACLE Live account setup
- RDBMS
- DQL Commands – select, where, group by, order by, having, distinct, limit, etc.
- DML Commands – insert, update, delete, lock, merge, etc.
- DDL Commands – create, drop, alter, truncate, rename, etc.
- DCL Commands – grant, revoke, deny, etc.
- TCL Commands – commit, save point, rollback, etc.
- Operators – Arithmetic operators, Comparison Operators, Logical Operators, String Operators
- Aggregated Functions – sum, avg, min, max, count, countd, pct, etc.
- Joins – Inner, Left, Right, Outer, Self, Cross
- Subquery – corelated subquery, nested subquery, etc.
- Analytical Functions – rank, dense_rank, row_number, etc.
- Views in SQL
- Date Functions – datepart, dateadd, extract, to_char, datetrunc, etc.
- Like Operator
- Advanced Query writing
- Built in functions for the database you are learning
- MS SQL SERVER Architecture
- MS SQL SERVER installation on local system
- MS SQL SERVER Commands and query writing practice (same topics as above)

MODULE 2: SQL QUERY WRITING

- 700 SQL Query writing practice questionary on different commands in w3 school.
- Hacker Rank Certification

Resources/ Study Links

Oracle DB - https://www.udemy.com/course/oracle-19c-sql/

MS SQL Server DB –
Udemy Course - udemy.com/course/the-complete-sql-course-2021-learn-by-doing/
YouTube - https://www.youtube.com/live/IjsfENuIJuY?feature=share

Theory resources for SQL –
w3 Schools - https://www.w3schools.com/sql/
javatpoint - https://www.javatpoint.com/sql-tutorial
geek for geeks - https://www.geeksforgeeks.org/sql-tutorial/

Query writing practice questions-
w3 resources – https://www.w3resource.com/sql-exercises/
Hacker Rank - https://www.hackerrank.com/domains/sql

TOPIC 5 – TABLEAU DESKTOP

Tableau Desktop is a powerful data visualization and business intelligence tool used by data analysts to analyse and present data in an interactive and meaningful way. As a data analyst, Tableau Desktop can greatly enhance your data exploration, analysis, and reporting capabilities

MODULE 1: TABLEAU INTERFACE

- What is tableau
- Architecture of tableau
- Features of tableau
- Installation of tableau desktop/public
- Interface of tableau (layout, toolbars, data pane, analytics pane etc)
- How to start with tableau
- Connect to data source
- Types of Data source
- Connect to Text File, Excel file, PDF, MS SQL Server, Spatial Files, etc.
- Introduction to the various file type
- How to create data visualization using tableau feature "show me"
- Working with workbook data and worksheet
- How to create calculated field
- Working with workbook data and worksheet
- Basic data visualization/graph
- Highlight table

MODULE 2: DATA Viz/ CHARTS/ GRAPHS

- Bar Chart, Stacked Bar Chart
- Bar in Bar Chart
- Combo Chart
- Line Chart, Area Chart
- Blended Axis
- Dual Axis Chart
- Lollipop Chart
- Pie Chart, Donut Chart
- Histogram Chart
- Animated Graph
- Map Chart and Custom Map Chart
- Pareto Chart
- Waterfall Chart
- Gant Chart
- Dumbbell Chart

TOPIC 5 – TABLEAU DESKTOP

- Sparklines
- Tiled and Hexagonal Maps
- Box Plots
- Explain latitude and longitude

MODULE 3: DATA PREPARATION

- Connecting to different Data Source
- Live vs Extract Connection
- Creating Extract
- Refreshing Extract
- Incrementing and Full Extract
- Refreshing Live
- Pivoting and splitting
- Data Interpreter: Clean Unstructured Data
- TWB vs TWBX
- How to create a packaged workbook
- Difference between .tde and .hyper file

MODULE 4: ADVANCED DATA PREPARATION

- Join Tables in Tableau- Inner, Left, Right, Outer
- Complex Join
- Referential Integrity
- Union
- Data Blending and Relationship Model
- Cross DB Join
- Custom Sequel Query writing

MODULE 5: FILTERING & MANAGING YOUR DATA

- Filter- Types of filter
- Action filter
- Global filter
- Normal filter
- Dimension filter
- Measure filter
- Condition based filter
- Advanced filter using wild card
- Right click filtering
- Top & Bottom N Filter
- Filtering Order of Operation

MODULE 6: BASIC & ADVANCED CALCULATIONS IN TABLEAU

- Date Functions
- String Functions
- Logical Functions
- Aggregate Functions
- Table Calculations
- User Functions
- LOD's- Fixed, Include, Exclude
- Window Functions
- Dynamic Calculations using Parameters

MODULE 7: DATA GROUPING

- Groups
- Sets
- Conditional sets, sets to find Top and Bottom
- Hierarchies
- Bins

MODULE 8: DASHBOARD & REPORTING

- Creating Dashboard
- Floating and Tiled Objects
- Sizing of Dashboard
- Dashboard Actions
- Container Lay outing
- Story making

MODULE 9: TABLEAU SERVER

- Installation of Tableau Online/ Cloud
- Publishing Dashboard
- Adding Users
- Setting Permissions to different users
- Roles on server
- Extract refresh
- Site status Dashboard
- Row level security (RLS)
- Alerts, subscriptions
- T links

MODULE 10: TABLEAU PROJECTS

- Project 1
- Project 2

Resources/ Study Links

Tableau Desktop -
Tableau Help (Theory concepts) - https://www.tableau.com/support/help
Tableau 2022 A-Z: https://www.udemy.com/course/tableau10/
Tableau 2022 Advanced: https://www.udemy.com/course/tableau10-advanced/
Tableau Calculations: https://www.youtube.com/watch?v=QimVQl5AoYM&list=PLkZ_g9YS
Forecasting and Time Series Analysis: https://www.udemy.com/course/forecasting-and-time-series-analysis-in-tableau/

Tableau Server -
Introduction: https://www.youtube.com/watch?v=zU8jetnJwz0
Intro to Tableau Server for Data Analyst: https://www.udemy.com/course/tableau-server-for-analysts/
Tableau Server: https://www.youtube.com/watch?v=d0-mB-0h-0g&list=PLWWK8OIhubE-LPLcxYeoNxyLeI0ww0VH-

Tableau Projects End to End-
https://youtu.be/oAIubTqg-Kw
https://youtu.be/pCpvlIr1yvs
https://youtu.be/NN5W_83N4Z4
https://youtu.be/NoppQVdd8U8
https://youtu.be/jhv_HPjtTyU

Tableau Material (Soft copy)-
1. Tableau hand written notes
2. Tableau printed notes
3. Tableau important formulae print notes
4. Tableau important interview questionnaire
All these are available at very minimal cost of Rs.200 only, ping me on WhatsApp number mentioned in description of this video.

TOPIC 6 – POWER BI DESKTOP

Power BI is a popular business intelligence and data visualization tool used by data analysts to analyse, visualize, and share insights from data. It offers a wide range of features and capabilities that support data exploration, analysis, and reporting.

MODULE 1: INTRO TO POWER BI

- Introduction to Power BI - Need, Importance
- Why and where Power BI is used?
- Power BI - Advantages and Scalable Options
- History - Power View, Power Query, Power Pivot
- Power BI Data Source Library and DW Files

MODULE 2: REPORT PROPERTIES & DIFFERENT ELEMENTS ON UI

- Report View Options: Full, Fit Page, Width Scale
- Report Design using Databases & Queries
- Power BI Design: Canvas, Visualizations and Fields
- Import Data Options with Power BI Model, Advantages
- General Properties, Sizing, Dimensions, and Positions
- Alternate Text and Tiles. Header (Column, Row) Properties
- Grid Properties (Vertical, Horizontal) and Styles

MODULE 3: DATA CONNECTIONS

- Connecting Power BI to flat files like excel, csv, pdf, etc
- Connecting Databases like MS SQL Server DB, My SQL, Oracle, PostgreSQL DB
- Connecting to Azure data
- Connecting to cloud services

MODULE 4: DATA MODELLING

- Connecting multiple tables to create a data model
- Relationship importance
- Understanding types of relationship
- Enable/Disable auto-relationships
- Create relationships manually
- Create relationship with cross database
- Marking as Active/Inactive relationship

MODULE 5: DATA VISUALIZATION

- Stacked bar chart, stacked column chart
- Clustered bar chart, clustered column chart
- 100% stacked bar chart, 100% stacked column chart
- Line charts, area charts, stacked area charts
- Line and stacked row charts
- Line and stacked column charts
- Waterfall chart, scatter chart, pie chart
- Field Properties: Axis, Legend, Value, Tooltip
- Field Properties: Colour Saturation, Filters Types
- Formats: Legend, Axis, Data Labels, Plot Area
- Data Labels: Visibility, Colour and Display Units
- Data Labels: Precision, Position, Text Options
- Analytics: Constant Line, Position, Labels
- Working with Waterfall Charts and Default Values
- Modifying Legends and Visual Filters - Options
- Map Reports: Working with Map Reports
- Format options to format all visuals and charts
- Conditional Formatting in charts

MODULE 6: HIERARCHIES / DRILLDOWN REPORTS

- Hierarchies and Drilldown Options
- Hierarchy Levels and Drill Modes - Usage
- Drill-thru Options with Tree Map and Pie Chart
- Higher Levels and Next Level Navigation Options
- Multi Field Aggregations and Hierarchies in Power BI
- DRILLDOWN, SHOWNEXTLEVEL, EXPANDTONEXTLEVEL

MODULE 7: POWER QUERY & M LANGUAGE

- Understanding Power Query Editor - Options
- Power BI Interface and Query / Dataset Edits
- Working with Empty Tables and Load / Edits
- Data cleaning
- Merge & Append tables
- Data source settings
- Transform data to Replace values, split columns, format columns, Reverse and Transpose, group by, etc.
- Add conditional columns, custom columns, column indexing
- Basics of M Query

MODULE 8: DAX EXPRESSIONS - BASIC

- Purpose of Data Analysis Expressions (DAX)
- Scope of Usage with DAX. Usability Options
- DAX Context: Row Context and Filter Context
- DAX Entities: Calculated Columns and Measures
- DAX Data Types: Numeric, Boolean, Variant, Currency
- Datetime Data Type with DAX. Comparison with Excel
- DAX Operators & Symbols. Usage. Operator Priority
- Parenthesis, Comparison, Arithmetic, Text, Logic
- DAX Functions and Types: Table Valued Functions
- Filter, Aggregation and Time Intelligence Functions
- Information Functions, Logical, Parent-Child Functions
- Statistical and Text Functions. Formulas and Queries
- Syntax Requirements with DAX. Differences with Excel
- Naming Conventions and DAX Format Representation
- Working with Special Characters in Table Names
- Attribute / Column Scope with DAX - Examples
- Measure / Column Scope with DAX - Examples

MODULE 9: DAX EXPRESSIONS - ADVANCED

- Aggregate Functions – SUM, SUMX, MAX, MAXX, MAXA, MIN, MINX, DIVIDE, PRODUCT, COUNT, DISTINCT COUNT, COUNTBLANK, AVG, etc.
- Filter Functions – ALL, ALLEXCEPT, ALLSELECTED, CALCULATE, FILTER, INDEX, LOOKUP, OFFSET, RANK, SELECTEDALUE, WINDOW, etc
- Date & Time Functions – CALENDAR, DATE, DATEDIFF, MONTH, YEAR, NOW, QUARTER, etc
- Logical Functions – AND, IF, OR, OR, SWITCH, NOT, FALSE, etc
- Information functions - CONTAINS, CONTAINSROW, ISAFTER, ISBLANK, ISEMPTY, ISERROR, etc
- Text Functions – FIND, LEFT, RIGHT, FIXED, MID, REPLACE, TRIM, UNICHAR, UNICODE, VALUE, FORMAT, CONCATENATE, EXACT
- Creating Date Table for Time Intelligence functions
- YTD, QTD, MTD Calculations with DAX
- DATESYTD, DATESQTD, DATESMTD
- ENDOFYEAR, ENDOFQUARTER, ENDOFMONTH
- FIRSTDATE, LASTDATE, DATESBETWEEN
- DAX Calculations and Measures
- Using TOPN, RANKX, RANK.EQ
- Computations using STDEV & VAR
- Data Analysis Expressions and Functions

- CLOSINGBALANCEYEAR, CLOSINGBALANCEQTR
- SAMEPERIOD and PREVIOUSMONTH, QUARTER
- Slicing and Dicing Options with Columns, Measures
- DAX for Query Extraction, Data Mashup Operations
- Calculated Columns and Calculated Measures with DAX

MODULE 10: POWER BI SERVICE

- Introduction
- Deployment (Publish to Report Server)
- Deployment (Uploading from Report Server)
- Deployment (Folder Structure) Working with Power BI Cloud Interface & Options
- Navigation Paths with "My Workspace" Screens
- Importing data from Desktop to Service
- Dashboard overview
- Workspace & Gateways
- Need for RLS in Power BI Cloud
- Row Level Security (RLS) with DAX

Power BI Desktop Basic Course -
https://www.youtube.com/watch?v=UXhGRVTndQA&t=951s
https://learn.microsoft.com/en-us/training/courses/pl-300t00

Power BI help for theory concepts - https://learn.microsoft.com/en-us/power-bi/

Power Query for Power BI
https://www.youtube.com/watch?v=VO-MyVMKUC4

Power BI Desktop Advanced Course -
https://www.udemy.com/course/microsoft-power-bi-up-running-with-power-bi-desktop/

DAX Expressions Course-
https://www.youtube.com/watch?v=yTOSOgUGKe4&t=1s
https://www.udemy.com/course/advanced-dax-for-power-bi/

Power BI Service-
https://www.youtube.com/watch?v=fmmQ4bhLfbo

Power BI Projects End to End-
https://youtu.be/-TErGczFxUs
https://youtu.be/Hn9f13uoLAQ
https://youtu.be/kfTVbJb9yIE
https://youtu.be/-sOHVl_iCHA

TOPIC 7 – PYTHON

Python is a popular programming language in the field of data analysis due to its versatility, rich ecosystem of libraries, and ease of use. It is used for Data Cleaning and Preparation, Exploratory Data Analysis (EDA), Statistical Analysis, Big Data Processing.

MODULE 1: INTRODUCTION TO PYTHON:

- Basic syntax and data types
- Variables and operators
- Control flow (if statements, loops)
- Functions and modules
- File I/O

MODULE 2: DATA MANIPULATION AND ANALYSIS:

- Introduction to NumPy and its array operations
- Introduction to Pandas and its data structures (Series, DataFrame)
- Loading and storing data in different formats (CSV, Excel, etc.)
- Data cleaning and pre-processing techniques
- Exploratory data analysis (EDA) with Pandas
- Handling missing data
- Data aggregation and summarization
- Merging, joining, and reshaping datasets

MODULE 3: DATA VISUALIZATION:

- Introduction to Matplotlib and basic plotting
- Creating line plots, bar plots, scatter plots, histograms, etc.
- Customizing plots with labels, titles, legends, etc.
- Introduction to Seaborn for advanced statistical visualization
- Interactive visualizations with libraries like Plotly or Bokeh

MODULE 4: Project Work:

- Application of Python and data analytics concepts to a real-world project
- Data cleaning, pre-processing, and analysis
- Visualization of insights
- Building and evaluating predictive models

Complete Python Bootcamp -
https://www.udemy.com/course/complete-python-bootcamp/
Data Analysis-
https://www.youtube.com/watch?v=obJZ1rB7TKc&t=129s
EDA-
https://www.youtube.com/watch?v=kLDTbavcmd0
Python Project-
https://www.youtube.com/watch?v=4QkYy1wANXA

TOPIC 8 – JIRA TOOL

JIRA is primarily known as a project management and issue tracking tool, but it can also be used in data analytics roles to streamline workflows, manage tasks, and track progress

- Introduction to JIRA - Overview of JIRA and its purpose in project management and issue tracking.
- User Interface and Navigation - Familiarization with the JIRA user interface and navigation.
- Issue Tracking and Management - Creating and managing issues (tasks, bugs, feature requests) in JIRA.
- Project Configuration - Creating and configuring JIRA projects.
- Agile Methodologies in JIRA - Introduction to Agile methodologies (Scrum, Kanban) and how they are implemented in JIRA.
- Difference between Agile and Waterfall Methodologies
- Tracking progress, managing dependencies, and visualizing work using Agile boards and reports.

Agile Methodology -
https://www.youtube.com/watch?v=h2Xzq2fbafM

JIRA Tool -
https://www.youtube.com/watch?v=Wgyirv9juFI

SCRUM CERTIFICATION
https://certiprof.com/pages/scrum-foundation-certificate-free

---

TOPIC 9 – SOFT SKILLS

In addition to technical skills, soft skills play a crucial role in the success of a data analyst. Developing and honing soft skills alongside technical skills will help data analysts excel in their roles and have a broader impact on their organizations.

- Communications skills – Verbal, Written, Presentation Skills
- Data Storytelling
- Attention to detail
- Curiosity and Learning Mindset
- Adaptability
- Cross Team Collaborations
- Curiosity and Learning Mindset

TOPIC 10 – RESUME, NAUKRI PROFILE & INTERVIEWS

- Resume Preparation
- Naukri Profile setup
- Mock Interviews

The expert in anything was once a beginner