

# Masters Data Analytics Training Curriculum

## STRUCTURE







#### Masters in Data Analytics

#### **About Croma Campus:**

Croma Campus Training & Development Private Limited is an Edtech Company since 2010 providing rigorous industry-relevant programs designed and delivered in collaboration with world-class faculty and industry with following features.

- Hands-On Live Projects
- Industry Cases Studies
- 61,000+ Satisfied Learners
- 140+ Training Courses
- 100% Certification Passing Rate
- Live Instructor Classroom / Online Training
- 100% Placement Assistance

#### Croma Campus Training Program Deliverables:

- Session Recordings Original Class Room Voice & Video Recording
- Training Material Soft Copy Handbooks
- Assignments | Multiple Hands-on Exercises
- **Test Papers** We provide **Practice Test** as part of our course to help you prepare for the actual certification exam.
- Live Case Studies
- Live Projects Hands-on exercises and Project work. You will work on real time industryoriented projects and assignments for each module to practice.
- Key focus on Hands-on exercises and Project work. You will work on real time industry-oriented projects.
- Faculty with more than **10+ Years of Experience** in the Industry.
- **Technical Resume Designing & Job Assistance:** With more than 100+ Clients across the Globe and we help learners to get a good job in their respective field. We also help learners with resume preparation.
- Interview Q&A
- About Croma Campus Training Certificate: Croma Campus will provide you with an industry-recognized (Certified by ISO 9001:2015 & E-Cell IIT Jodhpur) course completion certificate, which has lifelong validity.

How I unlock my Croma Campus Certificate: Attend Complete Batch & Submit at least One Completed Project.





#### Masters in Data Analytics Program Modules:

- Python for Data Analytics
- Data Analytics Overview
- > MS SQL For Data Analytics
- Advance Excel for Data Analytics
- PowerBI for Data Analytics
- Tableau for Data Analytics
- R Programming for Data Analytics
- Data Analytics Live Projects







## Course Content for Maters in Data Analytics:

## Module 1: Python Statistics for Data Science

#### • Introduction To Python:

- Installation and Working with Python
- Understanding Python variables
- Python basic Operators
- Understanding the Python blocks.

#### • Introduction To Variables:

- Variables, expression condition and function
- Global and Local Variables in Python
- Packing and Unpacking Arguments
- Type Casting in Python
- Byte objects vs. string in Python
- Variable Scope

#### • Python Data Type:

- Declaring and usingNumeric data types
- Using stringdata type and string operations
- Understanding Non-numeric data types
- Understanding the concept of Casting and Boolean.
- Strings
- List
- Tuples
- Dictionary
- Sets

#### • Introduction Keywords and Identifiers and Operators

- Python Keyword and Identifiers
- Python Comments, Multiline Comments.
- Python Indentation
- Understating the concepts of Operators

#### Data Structure

#### • List

- What is List.
- List Creation
- List Length
- List Append
- List Insert
- List Remove
- List Append & Extend using "+" and Keyword
- List Delete
- List related Keyword in Python
- List Revers
- List Sorting





- List having Multiple Reference
- String Split to create a List
- List Indexing
- List Slicing
- List count and Looping
- List Comprehension and Nested Comprehension

#### • Dictionary

- Dict Creation
- Dict Access (Accessing Dict Values)
- Dict Get Method
- Dict Add or Modify Elements
- Dict Copy
- Dict From Keys.
- Dict Items
- Dict Keys (Updating, Removing and Iterating)
- Dict Values
- Dict Comprehension
- Default Dictionaries
- Ordered Dictionaries
- Looping Dictionaries
- Dict useful methods (Pop, Pop Item, Str , Update etc.)

#### Sets, Tuples and Looping Programming

#### • Sets

- What is Set
- Set Creation
- Add element to a Set
- Remove elements from a Set
- PythonSet Operations
- Frozen Sets
- Tuple
  - What is Tuple
  - Tuple Creation
  - Accessing Elements in Tuple
  - Changinga Tuple
  - TupleDeletion
  - Tuple Count
  - Tuple Index
  - TupleMembership
  - TupleBuilt in Function (Length, Sort)
- Control Flow
  - Loops
  - Loops and Control Statements (Continue, Break and Pass).
  - Looping techniques in Python
  - How to use Range function in Loop?
  - Programs for printing Patterns in Python





- How to use if and else with Loop
- Use of Switch Function in Loop
- Elegant way of Python Iteration
- Generator in Python
- How to use nested IF and Else in Python
- How to use nested Loop in Python
- Use If and Else in for and While Loop
- Examples of Looping with Break and Continue Statements
- How to use IN or NOTkeywordin Python Loop.

### Exception and File Handling, Module, Function and Packages

#### • Python Exception Handling

- Python Errors and Built-in-Exceptions
- Exception handing Try, Except and Finally
- Catching Exceptions in Python
- Catching Specific Exception in Python
- Raising Exception
- Try and Finally

#### • Python File Handling

- Opening a File
- Python File Modes
- Closing File
- Writing to a File
- Reading from a File
- Renaming and Deleting Files in Python
- Python Directory and File Management
- List Directories and Files
- Making New Directory
- Changing Directory

#### • Python Function, Modules and Packages

- Python Syntax
- Function Call
- Return Statement
- Write an Empty Function in Python –pass statement.
- Lamda/ Anonymous Function
- \*argsand \*\*kwargs
- Help function in Python
- Scope and Life Time of Variable in Python Function
- Nested Loop in Python Function
- Recursive Function and Its Advantage and Disadvantage
- Organizing python codes using functions
- Organizing python projects into modules
- Importing own module as well as external modules
- Understanding Packages
- Programming using functions, modules & external packages





- Map, Filter and Reduce function with Lambda Function
- More example of Python Function

#### Data Automation (Excel, SQL, PDF etc)

#### • Python Object Oriented Programming—Oops

- Concept of Class, Object and Instances
- Constructor, Class attributes and Destructors
- Real time use of class in live projects
- Inheritance, Overlapping and Overloading operators
- Adding and retrieving dynamic attributes of classes
- Programming using Oops support

#### • Python Database Interaction

- SQL Database connection using
- Creating and searching tables
- Reading and Storing configinformation on database
- Programming using database connections

#### • Reading an excel

- Reading an excel file usingPython
- Writing toan excel sheet using Python
- Python | Reading an excel file
- Python | Writing an excel file
- Adjusting Rows and Column using Python
- ArithmeticOperation in Excel file.
- Plotting Pie Charts
- Plotting Area Charts
- Plotting Bar or Column Charts using Python.
- Plotting Doughnut Chartslusing Python.
- Consolidation f Excel File using Python
- Split of Excel File Using Python.
- Play with Workbook, Sheets and Cells in Excel using Python
- Creating and Removing Sheets
- Formatting the Excel File Data
- More example of Python Function

#### Working with PDF and MS Word using Python

- Extracting Text from PDFs
- Creating PDFs
- Copy Pages
- Split PDF
- Combining pages from many PDFs
- Rotating PDF's Pages

#### • Complete Understanding of OS Module of Python

- Check Dirs. (exist or not)
- How to split path and extension?
- How to get user profile detail?
- Get the path of Desktop, Documents, Downloads etc.





- Handle the File System Organization using OS
- How to get any files and folder's details using OS?

### Data Analysis & Visualization

#### • Pandas

- Read data from Excel File using Pandas More Plotting, Date Time Indexing and writing to files
- How to get record specific records Using Pandas Adding & Resetting Columns, Mapping with function
- Using the Excel File class to read multiple sheets More Mapping, Filling Nonvalue's
- Exploring the Data Plotting, Correlations, and Histograms
- Getting statistical information about the data Analysis Concepts, Handle the None Values
- Reading files with no header and skipping records Cumulative Sums and Value Counts, Ranking etc
- Reading a subset of columns Data Maintenance, Adding/Removing Cols and Rows
- Applying formulas on the columns Basic Grouping, Concepts of Aggregate Function
- Complete Understanding of Pivot Table Data Slicing using iLocand Locproperty (Setting Indices)
- Under sting the Properties of Pivot Table in Pandas Advanced Reading CSVs/HTML, Binning, Categorical Data
- Exporting the results to Excel Joins:
- Python | Pandas Data Frame Inner Join
- Under sting the properties of Data Frame Left Join (Left Outer Join)
- Indexing and Selecting Data with Pandas Right Join (Right Outer Join)
- Pandas | Merging, Joining and Concatenating Full Join (Full Outer Join)
- Pandas | Find Missing Data and Fill and Drop NA Appending DataFrameand Data
- Pandas | How to Group Data How to apply Lambda / Function on Data Frame
- Other Very Useful concepts of Pandas in Python Data Time Property in Pandas (More and More)

#### • NumPy

- Introduction to NumPy: Numerical Python
- Importing NumPy and Its Properties
- NumPy Arrays
- Creating an Array from a CSV
- Operations an Array from aCSV
- Operations with NumPy Arrays
- Two-Dimensional Array
- Selecting Elements from 1-D Array
- Selecting Elements from 2-D Array
- Logical Operation with Arrays
- Indexing NumPy elements using conditionals
- NumPy'sMean and Axis





- NumPy'sMode, Median and Sum Function
- NumPy'sSort Function and More

#### • MatPlotLib

- Bar Chart using Python MatPlotLib
- Column Chart using Python MatPlotLib
- Pie Chart using Python MatPlotLib
- Area Chart using Python MatPlotLib
- Scatter Plot Chart using Python MatPlotLib
- Play with Charts Properties Using MatPlotLib
- Export the Chart as Image
- Understanding plt. subplots () notation
- Legend Alignment of Chart using MatPlotLib
- Create Charts as Image
- Other Useful Properties of Charts.
- Complete Understanding of Histograms
- Plotting Different Charts, Labels, and Labels Alignment etc.

#### • Introduction to Seaborn

- Introduction to Seaborn
- Making a scatter plot with lists
- Making a count plot with a list
- Using Pandas with seaborn
- Tidy vs Untidy data
- Making a count plot with a Dataframe
- Adding a third variable with hue
- Hue and scattera plots
- Hue and count plots

#### • Visualizing Two Quantitative Variables

- Introduction to relational plots and subplots
- Creating subplots with col and row
- Customizing scatters plots
- Changing the size of scatter plot points
- Changing the style of scatter plot points
- Introduction to line plots
- Interpreting line plots
- Visualizing standard deviation with line plots
- Plotting subgroups in line plots

#### • Visualizing a Categorical and a Quantitative Variable

- Current plots and bar plots
- Count plots
- Bar plot with percentages
- Customizing bar plots
- Box plots
- Create and interpret a box plot
- Omitting outliers
- Adjusting the whiskers
- Point plots





- Customizing points plots
- Point plot with subgroups

#### • Customizing Seaborn Plots

- Changing plot style and colour
- Changing style and palette
- Changing the scale
- Using a custom palette
- Adding titles and labels: Part 1
- Face Grids vs. Axes Subplots
- Adding a title to a face Grid object
- Adding title and labels: Part 2
- Adding a title and axis labels
- Rotating x-tics labels
- Putting it all together
- Box plot with subgroups
- Bar plot with subgroups and subplots

## Module 2 : Data Analytics Overview

#### • Data Analytics Overview

- Dealing with Different Types of Data
- Data Visualization for Decision making
- Data Science, Data Analytics, and Machine Learning
- Data Science Methodology
- Data Analytics in Different Sectors
- Analytics Framework and Latest trends

## Module 3 : MS SQL For Data Analytics

#### • Introduction

- Overview of Oracle Database 11g and related products
- Overview of relational database management concepts and terminologies
- Introduction to SQL and its development environments
- The HR schema and the tables used in this course
- Oracle Database documentation and additional resources

#### • Retrieve Data using the SQL SELECT Statement

- List the capabilities of SQL SELECT statements
- Generate a report of data from the output of a basic SELECT statement
- Use arithmetic expressions and NULL values in the SELECT statement
- Invoke Column aliases
- Concatenation operator, literal character strings, alternative quote operator, and the DISTINCT keyword





Display the table structure using the DESCRIBE command

#### • Usage of Single-Row Functions to Customize Output

- List the differences between single row and multiple row functions
- Manipulate strings using character functions
- Manipulate numbers with the ROUND, TRUNC, and MOD functions
- Perform arithmetic with date data
- Manipulate dates with the DATE functions

#### • Conversion Functions and Conditional Expressions

- Describe implicit and explicit data type conversion
- Describe the TO\_CHAR, TO\_NUMBER, and TO\_DATE conversion functions
- Nesting multiple functions
- Apply the NVL, NULLIF, and COALESCE functions to data
- Usage of conditional IF THEN ELSE logic in a SELECT statement

#### • Aggregated Data Using the Group Functions

- Usage of the aggregation functions in SELECT statements to produce meaningful reports
- Describe the AVG, SUM, MIN, and MAX function
- How to handle Null Values in a group function?
- Divide the data in groups by using the GROUP BY clause
- Exclude groups of date by using the HAVING clause

#### • Display Data from Multiple Tables

- Write SELECT statements to access data from more than one table
- Join Tables Using SQL:1999 Syntax
- View data that does not meet a join condition by using outer joins
- Join a table to itself by using a self join
- Create Cross Joins

#### • Usage of Sub-queries to Solve Queries

- Use a Sub-query to Solve a Problem
- Single-Row Sub-queries
- Group Functions in a Sub-query
- Multiple-Row Sub-queries
- Use the ANY and ALL Operator in Multiple-Row Sub-queries
- Use the EXISTS Operator

#### • SET Operators

- Describe the SET operators
- Use a SET operator to combine multiple queries into a single query
- Describe the UNION, UNION ALL, INTERSECT, and MINUS Operators
- Use the ORDER BY Clause in Set Operations

#### • Data Manipulation

- Add New Rows to a Table
- Change the Data in a Table
- Use the DELETE and TRUNCATE Statements
- How to save and discard changes with the COMMIT and ROLLBACK statements
- Implement Read Consistency
- Describe the FOR UPDATE Clause





## Module 4 : Advance Excel for Analytics

#### • Ms Excel Basic

- Creation of Excel Sheet Data
- Range Name, Format Painter
- Conditional Formatting, Wrap Text, Merge & Centre
- Sort, Filter, Advance Filter
- Different type of Chart Creations
- Auditing, (Trace Precedents, Trace Dependents)Print Area
- Data Validations, Consolidate, Subtotal
- What if Analysis (Data Table, Goal Seek, Scenario)
- Solver, Freeze Panes
- Various Simple Functions in Excel(Sum, Average, Max, Min)
- Real Life Assignment work

#### • Ms Excel Advance

- Advance Data Sorting
- Multi-level sorting
- Restoring data to original order after performing sorting
- Sort by icons
- Sort by colours
- Lookup Functions
  - o Lookup
  - o VLookup
  - o HLookup
- Subtotal, Multi-Level Subtotal
  - Grouping Features
    - o Column Wise
    - o Row Wise
- Consolidation With Several Worksheets
- Filter

- o Auto Filter
- o Advance Filter
- Printing of Raw & Column Heading on Each Page
- Workbook Protection and Worksheet Protection
- Specified Range Protection in Worksheet
- Excel Data Analysis
  - o Goal Seek
  - o Scenario Manager
- Data Table
  - o Advance use of Data Tables in Excel
  - o Reporting and Information Representation
- Pivot Table
  - o Pivot Chat
  - o Slicer with Pivot Table & Chart





- Generating MIS Report In Excel
  - o Advance Functions of Excel
  - Math & Trig Functions
- Text Functions
- Lookup & Reference Function
- Logical Functions & Date and Time Functions
- Database Functions
- Statistical Functions
- Financial Functions
- Functions for Calculation Depreciation

#### • MIS Reporting & Dash Board

- Dashboard Background
- Dashboard Elements
- Interactive Dashboards
- Type of Reporting In India
  - Reporting Analyst
  - o Indian Print Media Reporting
- Audit Report
- Accounting MIS Reports
- HR Mis Reports
- MIS Report Preparation Supplier, Exporter
- Data Analysis
  - o Costing Budgeting Mis Reporting
  - o MIS Report For Manufacturing Company
  - MIS Reporting For Store And Billing
- Product Performance Report
- Member Performance Report
- Customer-Wise Sales Report
- Collections Report
- Channel Stock Report
- Prospect Analysis Report
- Calling Reports
- Expenses Report
- Stock Controller MIS Reporting
- Inventory Statement
- Payroll Report
- Salary Slip
- Loan Assumption Sheet
- Invoice Creation





## Module 5 : Power BI for Data Analytics.

#### • Introduction to Power BI

- Overview of BI concepts
- Why we need BI?
- Introduction to SSBI
- SSBI Tools
- Why Power BI?
- What is Power BI?
- Building Blocks of Power BI
- Getting started with Power BI Desktop
- Get Power BI Tools
- Introduction to Tools and Terminology
- Dashboard in Minutes
- Refreshing Power BI Service Data
- Interacting with your Dashboards
- Sharing Dashboards and Reports

#### • Power BI Desktop

- Power BI Desktop
- Extracting data from various sources
- Workspaces in Power BI
- Data Transformation
- Measures and Calculated Columns
- Query Editor

#### • Modelling with Power BI

- Introduction to Modelling
- Modelling Data
- Manage Data Relationship
- Optimize Data Models
- Cardinality and Cross Filtering
- Default Summarization & Sort by
- Creating Calculated Columns
- Creating Measures & Quick Measures

#### • Data Analysis Expressions (DAX)

- What is DAX?
- Data Types in DAX
- Calculation Types
- Syntax, Functions, Context Options
- DAX Functions
- Date and Time
- Time Intelligence
- Information
- Logical
- Mathematical
- Statistical





- Text and Aggregate
- Measures in DAX
- ROW Context and Filter Context in DAX
- Operators in DAX Real-time Usage
- Quick Measures in DAX Auto validations
- Power Pivot x Velocity & Vertipaq Store
- In-Memory Processing: DAX Performance

#### • Modelling with Power BI

- Introduction to Modelling
- Optimize Data Models
- Setup and Manage Relationships
- Cardinality and Cross Filtering
- Default Summarization & Sort by
- Creating Calculated Columns
- Creating Measures & Quick Measures

#### • Power BI Desktop Visualisations

- How to use Visual in Power BI?
- What Are Custom Visuals?
- Creating Visualisations and Colour Formatting
- Setting Sort Order
- Scatter & Bubble Charts & Play Axis
- Tooltips and Slicers, Timeline Slicers & Sync Slicers
- Cross Filtering and Highlighting
- Visual, Page and Report Level Filters
- Drill Down/Up
- Hierarchies and Reference/Constant Lines
- Tables, Matrices & Conditional Formatting
- KPI's, Cards & Gauges
- Map Visualizations
- Custom Visuals
- Managing and Arranging
- Drill through and Custom Report Themes
- Grouping and Binning and Selection Pane, Bookmarks & Buttons
- Data Binding and Power BI Report Server

#### • Introduction to Power BI Q&A and Data Insights

- Why Dashboard? and Dashboard vs Reports
- Creating Dashboards
- Configuring a Dashboard: Dashboard Tiles, Pinning Tiles
- Quick Insights in Power BI
- Power BI embedded and REST API
- Direct Connectivity
  - Custom Data Gateways
  - Exploring live connections to data with Power BI
  - Connecting directly to SQL Azure, HD Spark, and SQL Server Analysis Services/ My SQL
  - Introduction to Power BI Development API





- Excel with Power BI: Connect Excel to Power BI, Power BI Publisher for Excel
- Content packs
- Update content packs
- Publishing and Sharing
  - Introduction and Sharing Options Overview
  - Publish from Power BI Desktop and Publish to Web
  - Share Dashboard with Power BI Service
  - Workspaces and Apps (Power BI Pro) and Content Packs (Power BI Pro)
  - Print or Save as PDF and Row Level Security (Power BI Pro)
  - Export Data from a Visualization and Publishing for Mobile Apps

#### • Refreshing Datasets

- Understanding Data Refresh
- Personal Gateway (Power BI Pro and 64-bit Windows)
- Replacing a Dataset and Troubleshooting Refreshing

## Module 6 : Analytics with Tableau

#### • Introduction to Tableau2018

- What is Tableau?
- Features of Tableau
- Top Chart Types in Tableau
- Introduction to the various File Types
- Quick Introduction to the User Interface in Tableau
- How to Create Data Visualization Using Tableau feature "Show Me"
- Reorder & Remove Visualization Fields
- How to Sort & Filter Data
- How to Create a Calculated Field
- How to Perform Operations using Cross-Tab
- Working with Workbook Data & Worksheets
- How to Create a Packaged Workbook

#### • Tableau Architecture & User Interface

- Architecture of Tableau
- Installation of Tableau Desktop
- The interface of Tableau (Layout, Toolbars, Data Pane, Analytics Pane etc.)
- How to Start with Tableau?

#### • Data Preparation

- Connecting to Different Data Sources
- Excel
- CSV
- Microsoft Access
- SQL server
- Google Sheets
- Live vs. Extract Connection
- Creating Extract





- Refreshing Extract
- Incremental Extract
- Refreshing Live
- Data Source Editor
- Managing Metadata and Extracts
- Pivoting & Splitting
- Data Interpreter : Clean dirty data
- TWB vs. TWBX

#### • Data Visualization Principles

- What is Data Visualization?
- Why Visualization came into the picture?
- Importance of Visualizing Data
- Poor Visualizations versus Perfect Visualizations
- Principles of Visualizations
- Tufte's Graphical Integrity Rule
- Tufte's Principles for Analytical Design
- Visual Rhetoric
- Goal of Data Visualization
- Data Interpretation
- Pivot Tables
- Split Tables
- Responsive Tool Tips
- Radial & Lasso Selection
- Right Click Filtering
- Creating Calculated Fields
- Logical functions
- Case-if functions
- ZN function
- Else-if function
- Ad-Hoc Calculations
- Manipulating Text-Left and Right Functions

#### • Basic Data Visualization

- Pivot Table & Heat Map
- Highlight Table
- Bar Charts
- Line Charts
- Pie Chart
- Scatter Plot
- Word Cloud
- Tree Map
- Blended Axis
- Dual Axis

#### • Managing Your Data

- Filters
- Types of Filters
- Dimension Filters
- Measure Filters





- Condition based Filters
- Advanced filters using wildcards
- Top & Bottom N Filtering
- Filtering order of operations
- Extract Filter
- Data Source Filter
- Context Filter
- Other Filters etc
- Sorting
- Calculations String, Basic, Date & Logic
- Parameters
- Working with Dates
- Table Calculation
- Discrete vs Continuous measures
- Grouping Data
- Groups
- Sets
- Hierarchies
- Bins
- Combined Fields

#### • Formatting

- Size
- Updating Axis
- Colors
- Borders
- Transparency
- Chart Lines
- Trend Line
- Forecasting
- Reference Line
- Mark Labels
- Annotations

#### • Dashboard Design

- Canvas Selection & Adjusting Sizes
- Tiled Objects
- Floating Objects
- Pixel Perfect Alignment
- Summary Box
- Chart Titles & Captions
- Adding Images & Text
- Adding Background Color
- Adding Shading
- Adding Separator Lines
- Dynamic Chart Titles
- Information Icons
- Creating a Story
- Advanced Data Preparation





- Join
- Inner
- Left
- Right
- Full
- Complex Joins
- Union
- Data Blending & when it is required

## • Advance Data Visualization

- Bar Chart
- Stack Bar Chart
- Bar in Bar Chart
- Combo Chart
- Line Chart
- Single Axis
- Blended Axis
- Dual Axis
- Dual Axis Chart
- Line
- Bar
- Lollipop Chart
- Donut
- Pareto Chart
- Motion Charts
- Other Advanced Charts

## • Advanced Filtering & Actions

- Action Filters
- Action Jumps

## • Sharing Your Dashboards

- Publishing to PDF
- Exporting to Pivot Tables and Images
- Exporting Packaged Workbooks
- Publishing to Tableau Server

## Module 7 : Analytics with R Programming

- Overview
  - History of R
  - Advantages and disadvantages
  - Downloading and installing
  - How to find documentation?
- R Programming Basics
  - Using the R console and R Studio
  - Getting help





- Learning about the environment
- Writing and executing scripts
- Object oriented programming
- Introduction to vectorised calculations
- Introduction to data frames
- Installing and loading packages
- Working directory
- Saving your work

#### • Variable types and data structures in base R

- Variables and assignment
- Data types
- Numeric, character, Boolean, and factors
- Data structures
- Vectors, matrices, arrays, data frames, lists
- Indexing, sub-setting
- Assigning new values
- Viewing data and summaries
- Naming conventions
- Objects

#### • Getting data into the R environment

- Built-in data
- Reading data from structured text files
- Reading data using ODBC

#### • Data frame manipulation

- Introduction to tables, enhanced data frames
- Renaming columns
- Adding new columns
- Binning data (continuous to categorical)
- Combining categorical values
- Transforming variables
- Handling missing data
- Merging datasets together
- Stacking datasets together (concatenation)

#### • Handling dates in R

- Date and date-time classes in R
- Formatting dates for modeling

#### • Exploratory Data Analysis (Descriptive Statistics)

- Continuous data
- Distributions
- Quantiles, mean
- Bi-modal distributions
- Histograms, box-plots
- Categorical data
- Tables
- Bar plots
- Group by calculations





- Split-apply-combine
- Reshaping and pivoting data in R (long to wide with aggregation)
- Melt and cast

#### • Working with text data

- Finding and matching patterns in text
- Stringer package for text manipulation
- Introduction to regular expressions in R
- Categorical data wrangling with forcats

#### • Control flow & functions

- Truth testing
- Branching
- Looping
- Functions
- Parameters
- Return values
- Variable scope
- Exception handling
- Applying functions across dimensions
- Sapply, lapply, apply
- Programming with map and purr

#### • Graphics in R Overview

- Base graphics system in R
- Scatterplots, histograms, bar charts, box and whiskers, dot plots
- Labels, legends, titles, axes
- Exporting graphics to different formats

#### • Module 11: Advanced R graphics

- Understanding the grammar of graphics
- Quick plots (qplot function)
- Building graphics by pieces (ggplot function)
- Understanding geoms (geometries)
- Linking chart elements to variable values
- Controlling legends and axes
- Exporting graphics

#### • Inferential Statistics

- Bivariate correlation
- T-test and non-parametric equivalents
- Chi-squared test
- General Linear Regression Models in R
  - Understanding formulas
  - Linear and logistic regression models
  - Regression plots
  - Confounding / interaction in regression
  - Evaluating residuals
  - Scoring new data from models (prediction)
  - Useful plots from regression models