Specialisation Track: Insight Analyst

Point Estimation Ltd.

To be an excellent Insight Analyst, one must first understand the Business Objective of an Insight project. To extract meaningful insight, an Analyst must Understand the Data Structure, Run Exploratory Data Analysis, Prepare the Dataset adequate for analysis, Apply robust statistical methods to Derive actionable insights and Finally present it in an accessible way. The course is structured to cover all aspects of an Analyst role with hands-on exercise.

sourav.das@pointestimation.com ☎ 07948049226 14b Hogfair Lane, SL1 8BY, London, UK

COURSE SKILLSET

SQL | Actionable Insight Generation | Statistics One | Data Interpretation | Advanced Excel | Exploratory Data Analysis | Web Analytics | Google Analytics | Presentation | Python/R | Visualisation – DS, Tableau | Campaign Evaluation, A/B Testing | Intro to ML – Regression, Decision Tree, K-Means Clustering

SECTION 1: DATA PREP (INTRODUCTORY & ADVANCED SQL)

In sections 1 & 2 the focus will be on Data extraction and Manipulation techniques an Insight analyst must be adept at. We shall develop the same using SQL & Excel.

S1.1 Prep Work

- Installing SQL Server Management Service Attaching Database for the Module Exploring the Object Explorer
- Concept of Database and Tables

S1.2 Introduction to Data Structure

- How is Data Stored?
- What are the different Data Types? What are Records and Fields?

S1.3 Querying the Database-I

- Framework of the Query The SELECT statement
- Column Aliases - Distinct Function - Sorting the Output Dataset (Order By) Filtering the Data - Where Clause
- Wild Cards
- Adding commentary & Export Dataset to Excel

S1.4 Working with Tables

- Introduction on Objects & Schemas
- What are the different types of SQL Server Data Types? How to Create Tables and what are the Constraints?
- Column Properties - Indexing Columns
- Alter, Update, Insert, Delete, Drop, Truncate, Case Statements on Tables
- Adding, Altering, Deleting and Modifying Columns.

S1.5 Querying the Database - II

- Handling Null Data Datatype Conversion Aggregate Functions
- Introducing the concept of Group by Having
- Introduction to Statistical Functions Calculated Columns
- Concatenation
- Aggregate function with Over, Partition by Clause

S1.6 Querying the Database – III

- Concept of Venn Diagram - - - Exercise
- Joins: Inner Join, Outer Join, Left Join, Right Join, Self-Join, Cross Join

S1.7 Querying the Database – IV

- Concept of Nested/Sub-Queries

S1.8 Dealing with Dates

S1.9 Introduction to Database Normalisation

- What is normalisation? Relational DB design Various normal forms: 1NF, 2NF,3NF
- Pros and cons of normalisation

S1.10 Ranking and Value Window Functions

- Rank (), ROW_NUMBER (), DENSE_RANK (), PERCENT_RANK () CUME DIST(), NTILE()
- Lag (), Lead (), First_value (), Last_value () Handling Duplicates

S1.11 If Else Statement

- Testing single and multiple conditions using IF/ELSE constructs

S1.12 Looping

- While Loop

S1.13 Variables Declaring variables Setting Values

- Operations with variables Scope of variables
- Global Variables

S1.14 Stored Procedures

- Creating SP Altering SP
- Passing Parameters to a SP Introducing the concept of Group by Return Values

S1.15 Views

- Concept; how views work Creating and altering views

S1.16 Use of Temp Tables

- Advantages and Disadvantages of Temp Tables

S1.17 Merging

- Union, Union All, Merge

S1.18 Manipulating Strings & Advanced Functions

- Length (), Left (), Right (), Substring (), Trim() Character Indexing
- Replace Coalesce Cast, Convert
- Random Sampling using Rand Function

S1.19 Dynamic SQL S1.20 CTE

S1.21 Pivoting Data

S1.22 Automating/ Scheduling Queries / SP with Jobs S1.23 Using Import Export Wizard

S1.24 Assessment

- All concepts will be followed by hands on exercises
- There will be revisions, quiz and tasks in Google Classroom A final assignment and detailed assessment of SQL skills

SECTION 2: DATA PREP (INTRODUCTORY & ADVANCED EXCEL)

S2.1 Introduction

Excel Interface. Task Ribbon. Opening a blank/existing document. Workbook & Worksheets. Navigating through worksheets

S2.2 Working with Cells

Understanding & Formatting Cells. Merging Cells. Wrapping Cells. Adding Comments. Data Input. Rows & Columns. Inserting, Hiding rows/columns. Freezing Panes. Naming Cells and Ranges. Sorting Values.

S2.3 Working with formulas

Aggregation Functions . Introducing Statistical Functions. Copying Formulas using Relative and Absolute Cell Referencing. Introducing If Statements. Nested Statements. Error Trapping in formulas.

S2.4 Advanced Functions

AND/OR. SUMIF, AVERAGEIF, COUNTIF & corresponding IFS statements. Vlookup, Hlookup, Xlookup, Index, Small, Match. Date and Text Functions. Creating a Function. Working with Random number generators

S2.5 Understanding Data Structure & Tables

Tabular Structure. Inserting a table, converting a dataset into table. Table formatting. Filter and advanced filter. Sorting Data (Custom Sorting)

S2.6 Pivot Tables

Concept of Aggregating Data with Pivot Tables. Design table layout, grouping and ungrouping dimensions. Adding filters & Calculated columns. Expressing measure as sum, average, count, proportion etc. Using Get Pivot Data as cell reference.

S2.7 Creating Charts

Choosing appropriate Charts. Formatting charts - axes, chart titles, legends, major and minor axes. Swapping between rows and columns. Pivot Charts. Secondary axes. Automating Charts.

S2.8 Other Important Topics

Data Analysis Functionality. Statistical Analysis. Using solver. What-if analysis. Recording a macro. Conditional Formatting. Removing Duplicates. Connecting to SQL Databases. Introducing Power Pivot & DAX Measures. Shortcuts for Quick operation.

SECTION 3: EXPLORATORY DATA ANALYSIS (USING STATISTICAL METHODS)

In sections 3 & 4, we shall teach the fundamentals and techniques of Data interpretation – be it in a tabular structure or in forms of charts. We shall also be covering the best practices in Exploratory Data Analysis in this section. For example, before any data driven insight exercise, one must first understand the distribution, the consistency (with regards to the spread of the data points, etc.), the association, the influence an explanatory variable has over the response variable or on other explanatory variables. We shall train aspiring analysts on the fundamental concepts of Statistics that an Insight Analyst is expected to know to perform well in the role.

S3.1 Introduction

Sample Statistics vs. Population Parameter. Need for Statistical Inference. Variables and Data Types

S3.2 Descriptive Statistics

Central Tendency (Mean, Median, Mode). Dispersion (Standard Deviation, Variance, Coefficient of Variation). Range (Max, Min, Quartiles) Outlier Detection & Treatment. Distribution (Skewness and Kurtosis). Missing Variables and Imputation Techniques. Robust Statistics.

S3.3 - Data Distribution

Linear, Normal, exponential, logarithmic distributions etc. & Need for Transformation

S3.4 Hypothesis (Significance) Testing & Inferential Statistics

Z score. T-tests. Standard Error. Margin of Error. Critical Value. Significance level. Confidence Interval. One tailed and two tailed tests.

S3.5 Test of Association, Bi-Variate Statistics

Correlation. Co-variance. Chi-sq. Tests.

S3.6 ANOVA

Residuals. Degrees of Freedom. Mean Square Error. Sum of Squares (Total, Within, Between)

S3.7 Sampling Techniques

Random. Stratified. Clustered. Multistage, Systematic Random Sampling Techniques.

S3.8 Weighting Methods

Cell Weighting for Calibration/Tackling under or over representation. Light touch on other weighting methods.

SECTION 4: DATA INTERPRETATION

DI Fundamentals

S4.1 Introduction

The need for DI skills.

S4.2 Basic Algebra

Order of Operations (BODMAS rule). Manipulate Algebraic Expressions involving powers, logs, polynomials and fractions. Solving Simple Equations. Quadratic Equations. Inequalities. Indices. Logarithms and Exponential. Applying Summation and product.

S4.3 - Basic Probability and Set Theory

Set Theory. Venn Diagrams. Simple Probability. Conditional Probability. Probability of Independent Events.

S4.4 Interpreting Charts

Bar Charts. Column Charts. Stacked Column Charts. Scatterplots. Line Charts. Pie Charts. Combined Charts - Primary & Secondary Axes. Donut Charts. Radar Charts. Area Charts. 3D Charts.

S4.5 Profiling and Segmentation

The concept of Indexing (Over and Under-indexing attributes vs. Population distribution).

S4.6 Two Insight Projects

Exercise:

Project One. The first project will be a Sociology task trying to generate insights and establish relationship between an outcome variable and explanatory variables provided in the dataset. (Learnings: Correlation Analysis, Descriptive Statistics etc.)

Project Two. The second project will be on Customer Analytics - Profiling and Segmentation task using the techniques learnt in the above modules.

SECTION 5: CAMPAIGN DESIGN & EVALUATION

This section will cover all things about advertisement campaigns. We shall address the different types of campaigns (Call to Action, Brand Awareness, Retention, Acquisition Campaigns, Direct Response etc) run across media channels like TV (linear and VoD), Print, Digital media etc. In depth information will be given on how campaigns are executed: from choosing the objective and right target audience for the campaign to buying slots in TV or in digital space; from activating the campaigns to tracking the delivery; from measuring Campaign Delivery to finally evaluating the campaigns. End to end media planning process will be explained with a light touch on Market Mix Modelling.

S5.1 Introduction

Why are campaigns so essential to a Business?

S5.2 Different types of Campaigns

Brand Awareness. Retention. Acquisition. Call to Action. Direct Response. Engagement campaigns etc.

S5.3 Types of advertisement Inventory

On Linear TV (Slot, Sponsorship, Promos). On Digital Space (Banner, Display, Video, Social). On Print Media (Advertisement Columns). Programmatic Bidding.

S5.4 Audience Targeting

Campaigns are targeted differently. Linear TV Targeting Demographics. Digital Media Target Demographics. Print Media Target Demographics. eTelmar.

S5.5 Campaign Measures and Jargons

TVR. Impressions. Impacts. Views. Reach. Barb Glossary. Campaign delivery measurement & reporting (who,how & what). Digital Media Measures will be covered in the Web Analytics module.

S5.6 Campaign Evaluation

Measure uplift on Control vs Test Groups. Hypothesis Testing (same as A/B Testing in Digital and will be covered in Web Analytics). Calculate ROI.

S5.7 Media Planning & Effectiveness

Reducing Advertising Wastage. Optimise Spend, Channel Mix. Choosing the right target audience. Light Touch on Market Mix Modelling.

SECTION 6: WEB ANALYTICS

By the end of this module, you will have a complete understanding of Web Analytics terminologies. You will be comfortable in extracting Google Analytics data to gain insights on the traffic of a website. You will be able to create dashboards, measure performance against specific goals, build attribution model leading to a conversion and successfully perform A/B tests.

S6.1 Introduction

Web Analytics. E commerce. Google Analytics. Adobe Analytics.

S6.2 Web Analytics Glossary

Clicks. Serves. Impressions. Reach. SEO. PPC. Bounce Rate. CTR. VTR. Paid Search. Interstitial Ads. Frequency. Frequency Capping. Display. Banner Ads. Lead Generation. Attribution. Conversion. Conversion Rate, Goals... etc.

S6.3 Google Analytics Environment

Activating a Google Analytics Account. The Google Analytics Interface. Extracting Data to Excel.

S6.4 Google Analytics Reporting Metrics & Dimensions

A comprehensive list of all Dimensions & Measures and what do they mean.

S6.5 Google Analytics Dashboard

Dashboard Design. Key Traffic , Conversion, Engagement Insights.

S6.6 Google Campaign Manager

DCM. DSP. Tracking. Trafficking. Tag Management. Pixels. Campaign Report.

S6.7 Attribution Model

Last Touch. Multi Touch Attribution Models.

S6.8 CRO (Conversion Rate Optimisation)

Validating Web based CRO strategies using A/B Testing (One tailed and two tailed Hypothesis tests at different levels of significance)

S6.9 Google Analytics Certification

At the end of the module, you will appear the online test for GA certification.

SECTION 7: DATA VISUALISATION

Reporting is a key aspect of Analytics and to be able to communicate it well visually is extremely critical. This module in the specialisation track will highlight the Dos and Don'ts of Visual Analytics. There will be tutorials on how to use Tableau as a story board and sessions on Power BI. A session on Effective Insight Presentation (MS ppt) has been included in the module as practised by Blue Chip Companies.

S7.1 Introduction

The importance of visualisation in Analytics.

S7.2 More Charts

Continuation from Data Interpretation Module: Bollinger Bands. Bump Charts. Pareto Charts. Funnel Charts. Control Charts. Waterfall Charts.

S7.3 Tableau Tutorial

Tableau Interface Walk-through. Connecting to Database/Datasets. Filters. Parameters. Creating all the charts mentioned above and in the DI Module. Dashboard Designing. Custom SQL. Automate Reports. Share dashboard as HTML.

S7.4 Power BI

Connecting to Database/Datasets. Building Dashboard.

S7.5 Visualisation Check List

Dos & Don'ts of Visual Analytics.

S7.6 PowerPoint Presentation

Presentation Planner. Less Words, more figures. Key Insights as headlines.

S7.8 Assignment

Build a sales dashboard in Excel/Power BI/Tableau (sales database will be provided). Create a deck highlighting key insights from the data provided.

SECTION 8: LIGHT TOUCH MODELLING

S8.1 Introduction to Data Science

- Simple Linear Regression
- Model Assumptions. Model Adequacy Tests. Transformation. Model Evaluations.

S8.2 Classification Problem

- Decision Tree
- CHAID. Chi-Sq. Splitting Criterion. Tree Pruning. Node. Branches. Leaf.

S8.3 Cluster Analysis - Segmentation

- K-Means
- Euclidean Distance Measure, Convergence Criteria

S8.4 Hands-on Exercise

- Building a SLR & CHAID Model
- SLR Model will be used to predict house prices while a CHAID Model will be employed to create a rule set in distinguishing Churn Risky customer.
- Building a K-Means segmentation Model using Python.

SECTION 9: FINAL ASSIGNMENTS

S9.1 Insight Project

End to End project delivery of an Insight Project including Business Objective Gathering, Data Preparation, Exploratory Data Analysis Applying Statistical Methods, Producing a Dashboard/Report & A PowerPoint deck that needs to be presented back face to face. The Dataset will be provided on completion of all the modules.