SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

INTRODUCTION TO PROGRAMMING AND DATA

Introduction to programming and data

- Available Jobs for data analysis or data science.
- Understand What is after data analysis
- Understand Is learning programming in data analysis something complex or easy to learn it.
- Is Using MySQL give me final data simple for common audience
- Learn SQL versus Python versus R language.
- Could we imagine data science without SQL language?
- Understand Learning programming without prior experience.
- Learn What is big data really?
- Understand what is overloads in big data.
- Why SQL for data analysis
- **High-level overview about SQL projects.**
- (**Project 1**) Help Student in debt Crisis.
- (**Project 2**)Car crashing and worst drivers in some states in USA .
- (**Project 3**) DVD Rental store .
- Understand SQL theory basics.
- Learn Why SQL for data analysis
- Learn Relational databases
- Understand The logic behind SQL and SQL queries.
- Learn The full story between client and Dbms (database management system).

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

SQL FOR DATA ANALYSIS CONTENTS

MySQL for Data Analysis

- Learn How to start MySQL server.
- Learn How to stop MySQL server.
- Understand how MySQL IDE communicates with MySQL DB .
- Change database/schema name
- Learn Import external database to MySQL server .
- Understand MySQL data types.
- Learn Data Concatenate.
- Understand Create Tables.
- Understand the idea behind foreign key and primary key
- Learn Alter table .
- **Learn Modify one column .**
- **Learn Modify multiple columns .**
- **Learn Add one column .**
- **Learn Add multiple columns.**
- **Learn Rename column.**
- **Learn delete column.**
- Create REATIONAL tables (one to many)
- Understand Modify tables.
- Learn Use Cascade Action in relational tables.
- Learn Select statement
- Learn Limit statement
- Learn “From” statement .
- Learn “Where” clause .
- Understand Comparison
- Understand Logical Operators
- Understand Joins
- Understand Aggregate functions
- Learn Count function
- Learn Sum function
- Learn Min function
- Learn Max function

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

SQL FOR DATA ANALYSIS CONTENTS

MySQL for Data Analysis

- Learn AVG function
- Learn Date function
- Understand the difference between decimal and float.
- Understand Null values
- Understand Group by
- Learn Subqueries
- Understand Filtering
- Learn Order by clause
- Learn MySQL logic Operators
- Understand Equal operator
- Learn Greater than , less than operator
- Understand “Not equal” operator
- Understand “Between” operator versus “IN” operator.
- Understand “Like” operator.
- Learn “Is Null” operator
- Learn “Not” operator
- Understand All Join Clauses (inner join, outer join, full outer join)
- Understand **query vs sub-query .**
- Learn Import CSV file into MYSQL DB .
- Learn Export database to CSV file .
- Understand what is Comma separated value file (CSV) .

Bonus 1（Introduction to MongoDB）

- MongoDB for Data Analysis
- Understanding Non-relational database.
- Learn Why MongoDB for Data Analysis
- Learn What will happen if we used Relational tables in Big data ?
- Learn Languages used with MongoDB
- Understand Advantages and Disadvantages of Mongo DB
- Install MongoDB on Windows machine.
- Install MongoDB on MacOSX .
- **Write your first MongoDB query using JavaScript**

SYLLABUS

SQL & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

DATA ANALYSIS PROJECTS FOR BEGINNERS

Your First Data analysis project without Programming

- Learn how to answer questions using Excel chart
- Understand how to Analyze your problem into simple questions.
- Learn Filter and sorting table data using Excel table .
- Learn Generate final report as a Bar chart in Excel .
- Learn Generate different bar charts styles.
- Add Axis labels (X, Y label) to the bar chart .
- **Home work :** change bar chart bars colors .

Your Second Data analysis project without Programming

- Understand Statistical Analysis of the Work of Bob Ross
- Learn how to find average value for all table columns.
- Learn how to swap table column and rows .
- Learn to filter table values using “greater than” .
- Learn how to draw chart for table columns .
- Learn how to change bar chart styles
- **Home work :** add the following changes to the bar chart :
    - Change X axis values from fraction to percentage.
    - Add labels to X and Y axis .
    - Add main label title to chart .

SYLLABUS

SQL & **APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP**

INTRODUCTION TO STATISTICS FOR DATA SCIENCE

Applied statistics for data science

- Understand the difference between data analysis and data science.
- Understand Which courses and topics I should study it for data science.
- Understand What is big data?
- Understand overloads in big data.
- Understand volume, variety and velocity of data.
- Understand What jobs will be taken over by robots?
- Programming languages used with data analysis.
- **Data science jobs are the sexiest** jobs in the 21st century.
- Understand how professional people work in data analysis.
- Understand four main thing we should know them before study data analysis.
- Comparison between **inferential statistics and descriptive statistics.**
- Understand When to use descriptive statistics.
- Understand When to use inferential statistics.
- Data before and after descriptive statistics.
- Learn the topics I should focus on it in descriptive statistics.
- Understand population versus sample in inferential statistics.
- Understand why inferential statistics needs conclusion
- Understand why descriptive statistics do NOT need conclusions .
- Understand Why statistics in some cases could be a liar ?
- Understand what I need to learn before start data project.

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

CONTENTS OF DESCRIPTIVE STATISTICS

Descriptive statistics

- The importance of studying data types.
- Learn the two major types of data types.
- Understand categorical data.
- Understand numerical data.
- Data Types (Continuous vs. Discrete)
- Difference between Numerical and Categorical Data types
- Qualitative data and Quantitative data
- the main aspects for Quantitative Data
- main aspects of Categorical Data
- What are **Measures of Center?**
- what is mean?
- What is mode?
- What is median?
- Which one I should choose to measure center of the data?
- Measure center of the data helps us in our life society.
- What is spread
- Measures of spread
- Range
- Interquartile range
- Standard deviation
- Variance
- calculate Range and interquartile Range using a set of descriptive statistics
- Five number summary
- Why we need 5 number summary .
- Benefits of using Box Plot
- Comparison between histogram and boxplot
- Measures of Spread
- Range
- Inter-Quartile range
- Standard deviation
- Variance.

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

CONTENTS OF DESCRIPTIVE STATISTICS

Descriptive statistics

- Learn Calculate standard deviation using manual way.
- Learn Calculate standard deviation using automatic way.
- Understand the physical meaning of Standard deviation.
- Standard deviation versus center of the data.
- High standard deviation means high risk.
- Which one is better Range or standard deviation?
- Learn Data Shape
- Understand left skewed distribution
- Understand right skewed distribution
- Understand Symmetric distribution
- Understand Relationship between mean and median in normal Distribution.
- Understand Using Histogram and box plots for numerical data
- Box plot is very useful for 5 number summary
- Understand bell curve distribution
- Understand Gaussian distribution
- Understand Why normal distribution is very important ?
- Understand Difference between normal and Standard normal distribution
- Learn convert normal distribution to standard normal distribution
- Understand Why we need to convert to **standard** normal distribution?
- Outlier
- Understand Methods to find outliers
- Learn Use five number summary to determine if there is outliers
- What can i do with outliers ?
- What can I do if I removed outlier ?
- What can I do if I keep outlier ?

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

CONTENTS OF STATISTICS

Descriptive statistics (Normal distribution)

- Simplified viewpoint about Normal distribution
- Normal distribution is symmetric distribution , Why ?
- Understand How to convert normal distribution to standard normal distribution ?
- Understand Why we need to convert normal to standard normal distribution
- Is normal distribution used with z-statistic test .
- ***Normal distribution is the cousin of student "t" distribution.***
- Learn Z table.
- Understand standard normal distribution

Inferential statistics (Sampling distribution)

- Understand Population parameters
- Understand Sample statistics
- Difference between population parameters and sample statistics
- Understand measurement symbols for population parameters
- Understand measurement symbols for sample statistics
- Understand Why we need sample statistics?
- Use imaginary solution to find length of all fishes in the sea.
- Use sampling distribution to find mean length of population.
- Sampling distribution from a normally distributed population
- Understand the idea of sampling distribution.
- Understand The idea of central limit theorem.
- Learn The importance of confidence interval.
- Understand Confidence interval and Confidence level
- Learn **What do we mean by "confidence "?**
- Understand High width confidence interval
- Understand Low width confidence interval
- Understand What effects the width of Confidence interval
- Relationship between standard error and margin of error

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

CONTENTS OF INFERENTIAL STATISTICS

Inferential statistics (Sampling distribution)

- Understand the Difference between 99% confidence level and 95 % confidence level
- Comparison between High width confidence interval **and** low width confidence interval.
- Learn Effects the width of Confidence interval.
- relationship between standard error and margin of error
- understand relationship between confidence interval period and margin of error
- learn What is sample error?
- Learn What is success rate?
- Understand What is failures rate?
- Learn the difference between confidence level and success rate
- Understand What happens if we changed confidence level from 95% to 99% ?
- Make a decision based on High width or low width confidence interval
- Understand Trade-off between useless information and high risk in confidence interval
- Learn What effects the width of confidence interval?
- Understand Effect of standard deviation on confidence interval
- Understand Alpha? the proportion of confidence intervals that will **NOT** contain the population parameter

Inferential statistics (Student's T distribution)

- Introduction about student's t distribution
- Understand Why did we call t distribution by "student "?
- The founder of student's t distribution.
- Understand Comparison between t distribution and normal distribution
- Understand T distribution degree of freedom
- Learn What do we mean by degree of freedom?
- Understand The benefit of degree of freedom with t distribution
- Understand What will happen if DF > 30?
- Understand t table
- Understand t score and t table.

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

CONTENTS OF INFERENTIAL STATISTICS

Inferential statistics (student's t distribution)

- Understand t distribution and higher values of degree of freedom.
- Learn Calculate t value directly using t calculator.
- Understand when to use t or z distribution?
- Learn Confidence interval with t and z distribution
- Understand one tail Alpha and 2 tail Alpha
- Learn Student's t Distribution shape
- Learn Student's t distribution mean
- Examples and projects for t distribution and confidence level.
- Understand Margin of error with confidence interval 99%
- Understand Final results using t calculator
- Learn Two important equations for Z and T distributions
- Learn Use Excel to calculate confidence interval
- Learn using Excel to find confidence interval

Inferential statistics (HYPOTHESIS TESTING)

- Introduction about hypothesis testing
- Principles in hypothesis testing like H0 , H1 and P value
- Learn What is H0 (Null hypothesis) .
- Learn What is H1 (Alternative hypothesis) .
- Learn What is P value in hypothesis testing?
- P value is the most important thing we should focus on it
- Understand Analysis with H0 and H1 about Cola drink to see if the complaint is a fake
- Understand how to find H0 and H1 from Cola drinks
- On what basis you make decisions using H0 and H1
- Learn Alpha and type one error
- Comparing between Alpha and P value
- Understand P value and weak evidence.
- Understand P value and strong evidence.
- What is your decision for P value and weak evidence?

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

CONTENTS OF INFERENTIAL STATISTICS

Hypothesis testing (calculate P value)

- Learn the Steps you should follow up to make a decision.
- Learn how to Find H0 and H1 .
- Learn how to find Sample N,S, $\bar{x}$.
- Learn how to find Df (if we need it).
- Distribution type t or z.
- Learn how to find t or z value.
- Find P value from t table.
- Find P value from calculator directly.
- Comparing P and Alpha (α) value.
- **Make a decision**
- Use Excel to calculate P value.
- Prepare your excel sheet with sample results.
- Learn how to Write your P value equation in Excel file.
- **Make a decision by comparing Alpha and P value.**
- Understand one tailed hypothesis testing.
- Understand two tailed hypothesis testing.
- Understand difference between one and two tailed hypothesis testing

Hypothesis testing (two tailed t test)

- Understand the case of two tailed t test
- Two tailed test and Null hypothesis (**H0**) NOT equal to zero.
- Understand What is t critical?
- Understand What is Z critical?
- Understand One tail confidence versus two tail confidence
- Learn What is null hypothesis for two samples mean t test?
- Understand What is alternative hypothesis?
- Understand results from Excel file about two samples mean t test
- Use Excel to find t score for two tailed test.
- The critical value in one tail or two tail test.
- Statistical significance VS practical Significance
- Statistics is a liar or trustworthy

SYLLABUS

SQL – MYSQL FOR DATA ANALYSIS & APPLIED STATISTICS FOR DATA SCIENCE BOOTCAMP

CONTENTS OF INFERENTIAL STATISTICS

Bonus 2 (Machine learning)

- The Gentle introduction .
- Understand the difference between Machine learning And inferential Statistics.
- Understand What can we get from machine learning ?
- Understand prediction in machine learning .
- Learn Create decision tree method .
- Build your first machine learning method .
- Understand machine learning testing data set .
- Understand machine learning training data set .
- Understand Testing data set and the real data without prediction.
- Understand Why we need testing data set ?
- **Understand what we mean by**

  "predictions about the future" .