

uCertify

Exploratory Data Analysis with Python



Lesson



Practice test



Live-Lab

08 Jun 2023

- 1 Preface
- 2 Exploratory Data Analysis Fundamentals
- 3 Visual Aids for EDA
- 4 Activity: EDA with Personal Email
- 5 Data Transformation
- 6 Descriptive Statistics
- 7 Grouping Datasets
- 8 Correlation
- 9 Activity: Time Series Analysis
- 10 Hypothesis Testing and Regression
- 11 Model Development and Evaluation
- 12 Activity: EDA on Wine Quality Data Analysis
- 13 Appendix

1 

Get hands-on experience of Exploratory Data Analysis with Python with the comprehensive course and lab. The lab provides hands-on learning of EDA (Exploratory Data Analysis), beginning up with the basics to gain insights along with diverse techniques like data cleaning, data preparation, data exploration, and data visualization. The course and lab deal with importing, cleaning, and exploring data to perform preliminary analysis using powerful Python packages, and many more. Using Python for data analysis, you'll work with real-world datasets, understand data, summarize its characteristics, and visualize it for business intelligence.

2 3 

47

4 

63

5 

80

6 

80

7 

8 

9 

.

10 

- 2014
 - 1.
- 2015
 - 3.
- 2016
 - 3.
- 2017
 - 4.
- 2018
 - 3.
- 2019
 - 3.
- 2020
 - 3.

11 

1: Preface

- Who this course is for?
- What this course covers?
- To get the most out of this course

- Conventions used

2: Exploratory Data Analysis Fundamentals

- Understanding data science
- The significance of EDA
- Making sense of data
- Comparing EDA with classical and Bayesian analysis
- Software tools available for EDA
- Getting started with EDA
- Summary
- Further reading

3: Visual Aids for EDA

- Technical requirements
- Line chart
- Bar charts
- Scatter plot
- Area plot and stacked plot

- Pie chart
- Table chart
- Polar chart
- Histogram
- Lollipop chart
- Choosing the best chart
- Other libraries to explore
- Summary
- Further reading

4: Activity: EDA with Personal Email

- Technical requirements
- Loading the dataset
- Data transformation
- Data analysis
- Summary
- Further reading

5: Data Transformation

- Technical requirements
- Background
- Merging database-style dataframes
- Transformation techniques
- Benefits of data transformation
- Summary
- Further reading

6: Descriptive Statistics

- Technical requirements
- Understanding statistics
- Measures of central tendency
- Measures of dispersion
- Summary
- Further reading

7: Grouping Datasets

- Technical requirements
- Understanding groupby()
- Groupby mechanics
- Data aggregation
- Pivot tables and cross-tabulations
- Summary
- Further reading

8: Correlation

- Technical requirements

- Introducing correlation
- Types of analysis
- Discussing multivariate analysis using the Titanic dataset
- Outlining Simpson's paradox
- Correlation does not imply causation
- Summary
- Further reading

9: Activity: Time Series Analysis

- Technical requirements
- Understanding the time series dataset
- TSA with Open Power System Data
- Summary
- Further reading

10: Hypothesis Testing and Regression

- Hypothesis testing
- p-hacking

- Understanding regression
- Model development and evaluation
- Summary
- Further reading

11: Model Development and Evaluation

- Technical requirements
- Types of machine learning
- Understanding supervised learning
- Understanding unsupervised learning
- Understanding reinforcement learning
- Unified machine learning workflow
- Summary
- Further reading

12: Activity: EDA on Wine Quality Data Analysis

- Technical requirements
- Disclosing the wine quality dataset

- Analyzing red wine
- Analyzing white wine
- Model development and evaluation
- Summary
- Further reading

13: Appendix

- String manipulation
- Using pandas vectorized string functions
- Using regular expressions
- Further reading

35

PRE-ASSESSMENTS QUESTIONS

35

POST-ASSESSMENTS QUESTIONS

13 Live Labs

-

Exploratory Data Analysis Fundamentals

- Styling a Dataframe
- Applying Function to a Dataframe
- Slicing and Subsetting
- Dividing NumPy Arrays
- Inspecting NumPy Arrays
- Defining NumPy arrays
- Selecting rows
- Reading Data from a CSV File
- Creating a Dataframe

Visual Aids for EDA

- Creating a Line chart
- Creating a Bar Chart
- Creating a Scatter Plot

- Creating a Bubble Chart
- Creating an Area Plot
- Creating a Pie Chart
- Creating a Table Chart
- Creating a Polar Chart
- Adding the Best-Fit Line for the Normal Distribution
- Creating a Histogram
- Creating a Lollipop Chart

Activity: EDA with Personal Email

- Performing EDA with Email Data
- Extracting Email Using Regex
- Converting a Field to datetime
- Removing NaN Values
- Dropping a Column

Data Transformation

- Stacking a Dataframe
- Concatenating Dataframes
- Analyzing Dataframes
- Combining Dataframes
- Merging on Index
- Permuting a Dataframe
- Removing Duplicate Data
- Replacing Values
- Interpolating Missing Values
- Backward and Forward Filling
- Handling NaN values
- Counting Missing Values
- Renaming Axis Indexes
- Binning
- Detecting Outliers

Descriptive Statistics

- Generating a Binomial Distribution Plot
- Generating an Exponential Distribution Plot
- Generating a Normal Distribution Plot
- Generating a Uniform Distribution Plot
- Using Statistical Functions
- Calculating Standard Deviation
- Finding Skewness and Kurtosis
- Creating a Box Plot
- Calculating Inter-Quartile Range

Grouping Datasets

- Finding Maximum Value for Each Group
- Grouping a Dataset
- Filtering Data
- Applying Aggregation Functions
- Creating a Pivot Table
- Creating a Cross-Tabulation Table

Correlation

- Calculating Correlation Coefficient

Activity: Time Series Analysis

- Sampling the Data
- Resampling the Data
- Changing the Index of a Dataframe

Hypothesis Testing and Regression

- Performing Z-Test
- Calculating the P-Value

- Performing T-test
- Scoring the Model
- Understanding the Linear Regression Model

Model Development and Evaluation

- Using TfidfVectorizer

Activity: EDA on Wine Quality Data Analysis

- Plotting a Heatmap
- Visualizing the Data in 3D Form

Appendix

- Accessing Characters
- String Slicing
- Updating a String
- Escape Sequencing
- Formatting Strings
- Displaying Last 10 items from a Dataframe
- Using String Functions with a Dataframe
- Finding Words from a String
- Counting Full Stops using Regex
- Matching Characters

77
LIVE LABS

13
VIDEO TUTORIALS

20
MINUTES

14 



support@ucertify.com

