Statistics for Data Science with Python (coursera_ST0151EN)

Week 1_Course Introduction and Python Basics and Descriptive Statistics

This module will focus on introducing the basics of descriptive statistics - mean, median, mode, variance, and standard deviation. It will explain the usefulness of the measures of central tendency and dispersion for different levels of measurement.

Learning Objectives

- **Course Introduction and Python Basics**
    - Introduction to the course and Instructors
    - Basics of Python and Jupyter Notebooks
- **Introduction & Descriptive Statistics**
    - Compute the measures of central tendency and dispersion
    - Understand the usefulness of the measures of central tendency and dispersion
    - Understand types of data

Week 2_Data Visualization

This module will focus on different types of visualization depending on the type of data and information we are trying to communicate. You will learn to calculate and interpret these measures and graphs.

Learning Objectives

- Interpret graphical summaries of data
- Create data visualizations in Python
- Effectively choose the right chart type for the audience and data type

Week 3_Introduction to Probability Distributions

This module will introduce the basic concepts and application of probability and probability distributions.

Learning Objectives

- Calculate probabilities given a Normal density
- State your null and alternative hypothesis when performing tests
- Understand the different bell-shaped distributions

Week 4_Hypothesis Testing

This module will focus on teaching the appropriate test to use when dealing with data and relationships between them. It will explain the assumptions of each test and the appropriate language when interpreting the results of a hypothesis test.

Learning Objectives

- Conduct a hypothesis test on a population mean
- Correctly formulate a decision rule for testing a hypothesis
- Conduct a hypothesis test on a difference in two or more population means
- Distinguish between correlation tests for two continuous variables and two categorical variables

总结:

Week 5_Regression Analysis

This module will dive straight into using python to run regression analysis for testing relationships and differences in sample and population means rather than the classical hypothesis testing and how to interpret them.

Learning Objectives

- Conduct correlation tests using a simple linear regression
- Conduct a ANOVA using a simple linear regression
- Conduct a t-test using a simple linear regression

Week 6_Project Case: Boston Housing Data

In the final week of the course, you will be given a dataset and a scenario where you will use descriptive statistics and hypothesis testing to give some insights about the data you were provided. You will use Watson studio for your analysis and upload your notebook for a peer review and will also review a peer's project. The readings in this module contain the complete information you need.

Learning Objectives

- Use Watson Studio
- Create a Jupyter Notebook
- Upload your Notebook
- Review a peer's notebook

总结: