Google Advanced Data Analytics Capstone

Instructions

Use this PACE strategy document to record your decisions and reflections as a data professional as you work through the capstone project. As a reminder, this document is a resource guide that you can reference in the future and a space to help guide your responses and reflections posed at various points throughout the project.

Portfolio Project Recap

Many of the goals you accomplished in your individual course portfolio projects are incorporated into the Advanced Data Analytics capstone project including:

- Create a project proposal
- Demonstrate understanding of the form and function of Python
- Show how data professionals leverage Python to load, explore, extract, and organize information through custom functions
- Demonstrate understanding of how to organize and analyze a dataset to find the "story"
- Create a Jupyter notebook for exploratory data analysis (EDA)
- Create visualization(s) using Tableau
- Use Python to compute descriptive statistics and conduct a hypothesis test
- Build a multiple linear regression model with ANOVA testing
- Evaluate the model
- Demonstrate the ability to use a notebook environment to create a series of machine learning models on a dataset to solve a problem
- Articulate findings in an executive summary for external stakeholders

Project proposal

Salifort Motors: Increase in Employee Retention and Satisfaction

Overview

Analyze Employee Survey responses and create a model that can help predict whether an employee will leave the company based on the employee survey data .

Milestones	Tasks	PACE stages
1 Week	- Gather as much data as possible pertaining employees and relevant information - Conduct Preliminary analysis of gathered data to begin brainstorming approach to the business problem (Employee Satisfaction and Retention) - Prep Data for EDA	**Plan**
1-2 Weeks	- **Perform Exploratory Data analysis by addressing missing values, duplicates, incorrect data types and potentially doing some feature engineering.** - **Gather Basic Statistical Data to help understand raw data and lead to pattern discoveries.** - **Create Basic Visuals to present to Stakeholders** - **Meet with Stakeholders to share findings**	Analyze
1 Week	- **Reflect on Preliminary Analysis** - **Discuss Potential Model/s** - **Prepare data for model creation and rollout**	Plan
2-3 Weeks	- **Choose and proceed with creating initial model/s** - **Document findings** - **Create Stakeholder Summary**	Construct
1-2 Weeks	- **Refine Model as needed by stakeholders** - **Document New Findings** - **Create Stakeholder Summary**	Construct 2.0
1-2 Months	- **Stakeholders and Company Roll out Model**	Execute