Project Proposal

This repo is one of the T5 Bootcamp requirements.

Forecasting of Employee Salary

It is a model enable Stakeholder like MOF, Sponsors, companies, Business owners, and project managers to predict what they will spend on their employees based on their years of experience.

Dataset

Using dataset from Kaggle for 4 years (2011,2012, 2013, and 2014) it simulated the confidential data frame that I have.

Dataset include 148654 records and 13 features

as fellow:

Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay
	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency
	Status					

Features which contains linking data to a more clearly visible picture, are used to identify the most important features of this study.

Tools

There are tools that will be used to achieve the goal of this study, such as: `matplotlib, pandas, seaborn, sklearn` for discovering the data and train a model. The work will be done through Jupyter notebook.

TO DO:

- Explore the data and come up with EDA phases then use a model to fit the data.