Assignment-4

Q1) The median salary of male employees is

A screenshot of a social media post

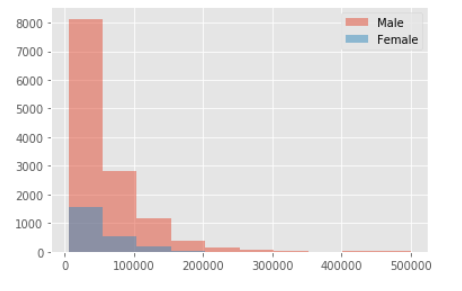
Description automatically generated

The median salary of female employees is

A screenshot of a social media post

Description automatically generated

Q2) Histogram graph for the salaries of male and female employees



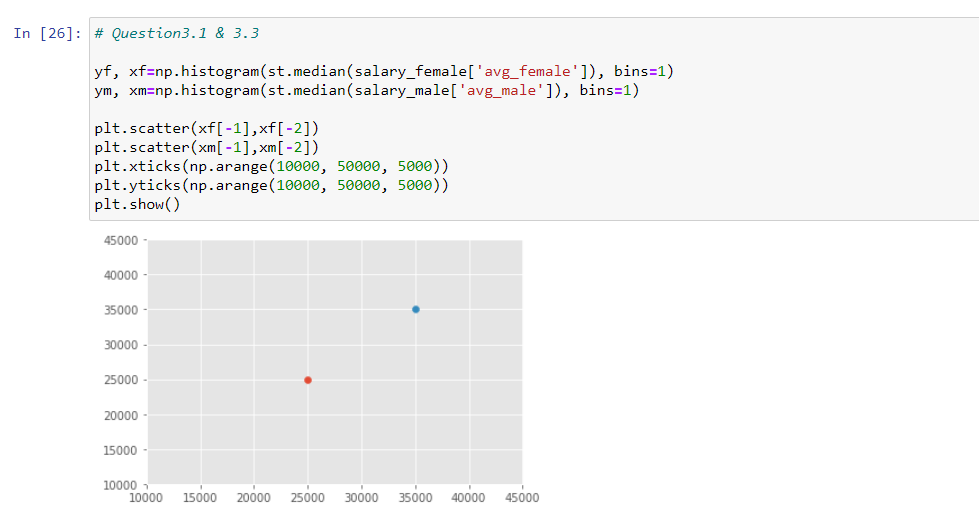
Q3)

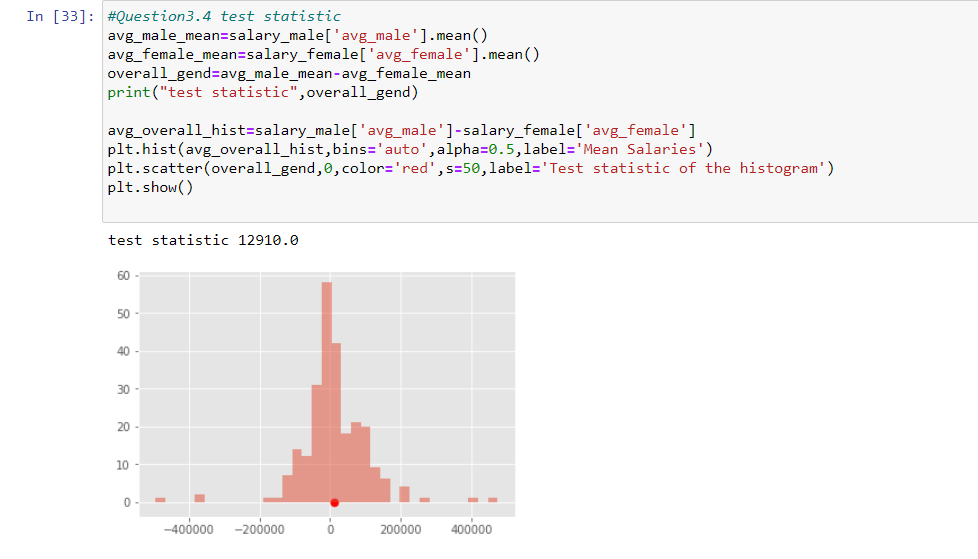
**Test Hypothesis**: The numerical description of the dataset that reduces data to one value that can be used for the testing hypothesis.

**Null Hypothesis**: No statistical difference between two variables then it is defined as null hypothesis.

**Alternate Hypothesis**: It is defined as there is a significant difference between two variables.

Q 3.1 & 3.3)



Q4) 

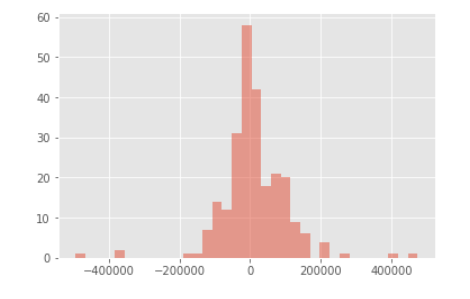
Q3.5)

**Bootstrapping:** It is a computational method that creates a large number of simulated samples from a single dataset. This is used to calculate the standard errors, confidence intervals and hypothesis testing and sample statistics such as mean, median and standard deviation can be determined from single large sample.

**Procedure:**

1. The bootstrap approach has an equal probability of randomly creating 5000 samples from the original data.
2. For the resampled dataset, the property “with replacement” is used for picking one data point more than once.
3. The method produces resampled datasets equal to the original dataset.
4. This method uses sample statistic distribution across virtual samples as the distribution of samples.

Q 3.6)



Q 3.7)

A confidence interval is determined from the statistics of the observed data.

The P-value is the marginal significant point within a statistical hypothesis test

