Hypothesis testing is a statistical method to test a statement from the research group which tend to reject or accept the null values based on the results from the experiment conducted using the random samples generated from the original. We generally use this type of hypothesis testing when we have very low amount of data to work on the statistical analysis of a data. We initially have the null hypothesis H0 which is the truth value or initial value of the statement from the data we have and another is alternate hypothesis which might be mostly the opposite of the null hypothesis.

I would like to work on the **Representation of women in low-paying jobs** from the list of the items in the canvas. The hypothesis for this statement would be

H0 : - women with low paying > 50% are overrepresented because of the population difference between men and women.

HA: - Women with low paying > 50% are not overrepresented and it’s a fact that women are earning less than men.

From the chart the women are considered as low pay when their hourly wages are less than 2/3rd of the median of the hourly earnings of the person (Irrespective of men or women). When the percentage of the low wages for women goes more that 50% it implies that they are overrepresented as low wage earners, but it happens because of the overpopulation of women over the men and it seems to be true when we check compare it with population of women in the country **“Switzerland”.** Switzerland has more women than men and obviously this might be the reason to get more percentage of low wage workers on the women than the men. But this might not be the true for all the countries, to check that we need to compare it with the percentages of the women all over the world to the countries ration and then decide on what percent of the low wage earners are from the women and men.

In the above scenario the significance level can be set to 5% and then carry out the test statistics and the corresponding p-value on the other hand comparing it with real time values is much needed to maintain a good accuracy of the model, as the P-values tend to change based on the size of the random samples generated which are used for testing the hypothesis. We use the confidence interval as the deciding factor on the rejecting the null hypothesis as there might be sample that have a very less relationship between the variables in the data of the random samples generated for testing the hypothesis.

We need to reject or accept the null hypothesis based on the conditions like P-value <= significance level then we can reject our null hypothesis and P-value > significance level we fail to reject the null hypothesis which means the initial statement made by the researches is not correct according to the hypothesis testing.