The Difference in monthly income by gender in the company

The gender pay gap is the difference between men's and women's average earnings from employment, shown as a percentage of men's average earnings.

The UNECE gender statistics database shows that this is one of the most significant problems of inequality between genders. It is from this problem that our study is designed to get rid of the example of a particular company. We explored a dataset with complete information about employees.

In our study, we want to look at the income gap in a particular company by gender to improve the environment and promote equality between genders.

It is also necessary to understand how salaries are distributed depending on age, whether there is a correlation between other indicators affecting wages, and how the salary budget is distributed between different departments.

Also, we must:

- Prepare the model in Power BI using the .CSV file provided
- Make the dataset into a star schema
- Create measures using variables
- Identify outlier areas

First, we met with our dataset. The data set consists of thirty-five columns, and 1470 rows there are no gaps in the data and duplicated rows. We are interested in columns with information on gender, monthly income, and age.

And to normalize the data, we made a model in POWER BI. Our model consists of one "fact" table named "employee_fact" and nine-dimension tables with confidential information about employees, job responsibilities, job experience, employee education, income information, job position, job survey (questioning about work satisfaction), and exploratory data.

All dimensions tables related to the fact table with the relation "many to one" with a key column, except exploratory data, which is related to "one to one." The model is accepted as the star schema.

For the next step, we tried to find outliers, for this reason, we called the describe method in Python, there we found too big salary of 19999, which was a maximum and with a significant difference, for example, the third quartile was 8379. We did the bow plot. The graph shows that anything over 17000 is an outlier and has a strong effect on the average.

We did the measure, which shows the ratio between male and female monthly income. In common, this ratio equals 95,45%, which means that our monthly income for female persons is quite larger than for male persons.

In our company 40% of Female workers and 60% The most. Most equal salary in the Sales department, their ratio is 99,8%.

We tried to find statistical differences between the average income of male and female workers.

We used the Mann-Whitney test to compare means, since we cannot speak of a normal distribution for income levels, and Student's test is not suitable for us.

The Hypothesis testing about comparing the means of two control groups:

H_0: Mean income in groups is equal

H_a: Mean income in groups is different
alpha = 0.05

After testing we got the next results: p-value: 0.08841668326602112, Cannot reject the null hypothesis, and a conclusion about the difference cannot be made.

That shows that there is no discrimination in our company related to gender.

Also, we considered the distribution of the company's salary budget among employees, by gender. We showed that the budget spent on the salary of men is larger, which is not surprising since there are more men in the company.

We checked the correlation between monthly income and other numeric data. According to the Chaddock scale: there is a weak linear relationship between wages and years in the current position. An average linear direct relationship is observed between the level of salary and age, as well as how many employees work in this company. A high direct linear relationship is found between the level of salary and overall work experience, and an extremely high relationship, which is logical, between salary and the level of work.

Also, we can see it in the graphs.

Conclusions.

1. We checked the data of staff on gender equality income.
2. In our research we show, that in this company there is no significant difference between the average income of male and female workers.
3. Moreover, it is shown that in some departments, women's salaries are even higher
4. Work experience is the main factor on which salary depends.
5. A larger budget for men's salaries is since there are more men in the company in the ratio of 40% women, 60% men. HRs should pay attention to the unequal occupation of positions.