

Relationships between Continuous Variables Lab

Load Glassdoor: <https://github.com/fenago/cts245X/tree/main/EDA>

Part 1: Analyzing relationships with scatter plots

On a new page, named "scatter plots", create a new scatter plot of `YearsAtCompany` vs. `MonthlyIncome`.

Create a new scatter plot of `TotalWorkingYears` vs. `MonthlyIncome`.

Create a new scatter plot of `JobLevel` vs. `MonthlyIncome`.

Create a new scatter plot of `NumCompaniesWorked` vs. `MonthlyIncome`.

Which variable has the strongest positive relationship with `MonthlyIncome`?

- ☐ `JobLevel`
- ☐ `YearsAtCompany`
- ☐ `TotalWorkingYears`
- ☐ `NumCompaniesWorked`

Part 2: Adding context to scatter plots

On a new page, called "scatter plot color", copy the previous exercise's scatter plot showing `MonthlyIncome` vs. `JobLevel`.

Add color to the story of the scatter plot by adding `JobRole` to the *Legend*. Place the legend on the bottom of the chart.

Format the scatter plot to maintain a consistent x-axis and y-axis. In order to do this, make a change in the ranges of the axes.

Create a *Stacked bar chart* showing the count of reviews by `JobRole`.

Finally, create two *Card* visualizations - one for the median `MonthlyIncome` and the other for the average `MonthlyIncome`. Arrange the cards to be at the top of the page and the charts on the bottom.

Using the bar chart to filter the other visualizations by `JobRole`, explore the relationship between `JobLevel` and `MonthlyIncome` changes for each separately.

Which job role has the highest median monthly income?

Part 3: Correlation coefficient

Add a trend line to the scatter plot to show the general relationship between `JobLevel` and `MonthlyIncome`.

Calculate the correlation coefficient using a "New quick measure".

Create a new *Card* visualization for the correlation coefficient. Arrange the cards to be evenly spaced on top.

What is the correlation coefficient for the relationship between `JobLevel` and `MonthlyIncome`?