```r
jobs_in_data1<-read.csv("//Users//chaithanayayennam//Downloads//jobs_in_data.csv")
head(jobs_in_data1)
```

```
##   work_year            job_title                  job_category salary_currency
## 1      2023 Data DevOps Engineer              Data Engineering             EUR
## 2      2023       Data Architect Data Architecture and Modeling             USD
## 3      2023       Data Architect Data Architecture and Modeling             USD
## 4      2023       Data Scientist      Data Science and Research             USD
## 5      2023       Data Scientist      Data Science and Research             USD
## 6      2023       Data Scientist      Data Science and Research             USD
##   salary salary_in_usd employee_residence experience_level employment_type
## 1  88000         95012            Germany        Mid-level       Full-time
## 2 186000        186000      United States           Senior       Full-time
## 3  81800         81800      United States           Senior       Full-time
## 4 212000        212000      United States           Senior       Full-time
## 5  93300         93300      United States           Senior       Full-time
## 6 130000        130000      United States           Senior       Full-time
##   work_setting company_location company_size
## 1       Hybrid          Germany            L
## 2    In-person    United States            M
## 3    In-person    United States            M
## 4    In-person    United States            M
## 5    In-person    United States            M
## 6       Remote    United States            M
```

```r
# Descriptive statistics for quantitative variable "Salary"
summary(jobs_in_data1$salary)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   14000  105200  143860  149928  187000  450000
```

```r
# Descriptive statistics for categorical variable "Job Title"
table(head(jobs_in_data1$job_title))
```

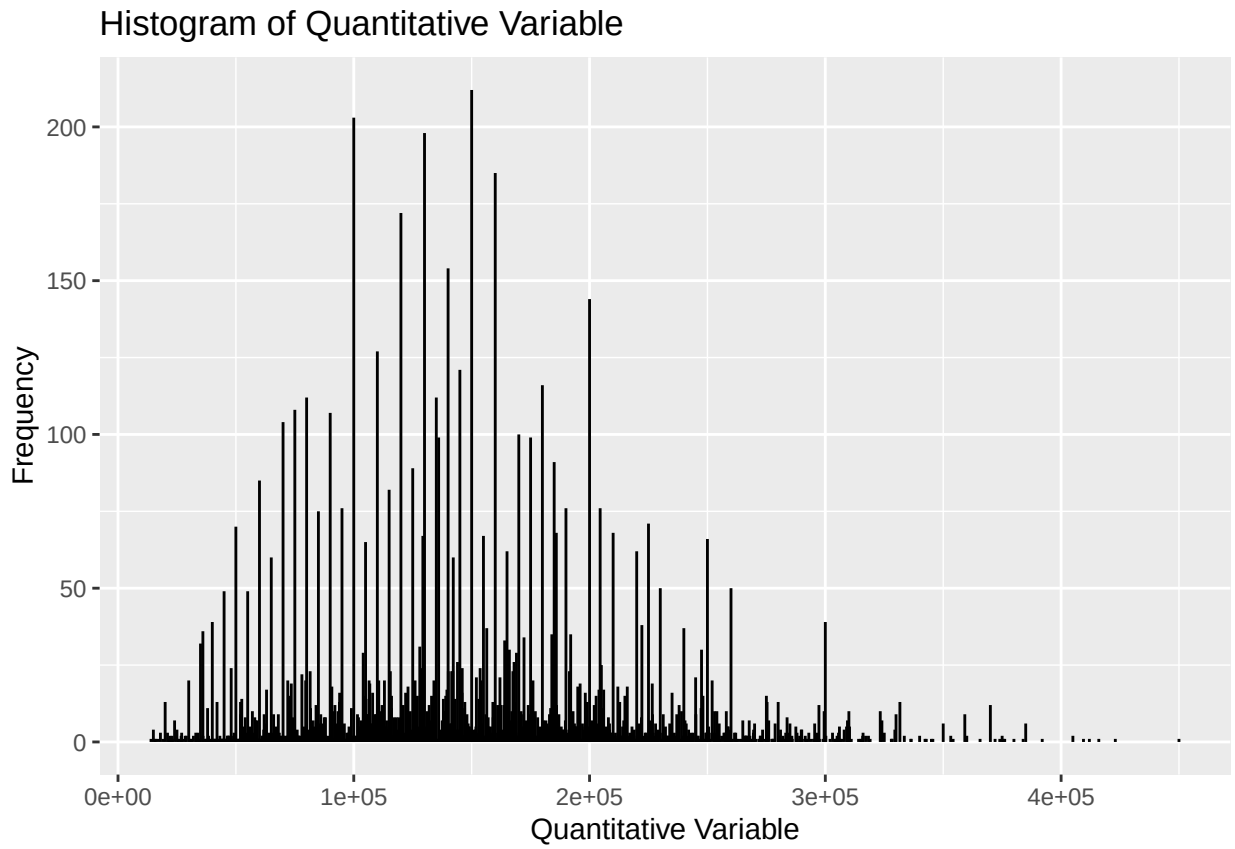
```
##
##       Data Architect Data DevOps Engineer       Data Scientist
##                    2                    1                    3
```

```r
jobs_in_data1$TransformedSalary <- sqrt(jobs_in_data1$salary)
head(jobs_in_data1)
```

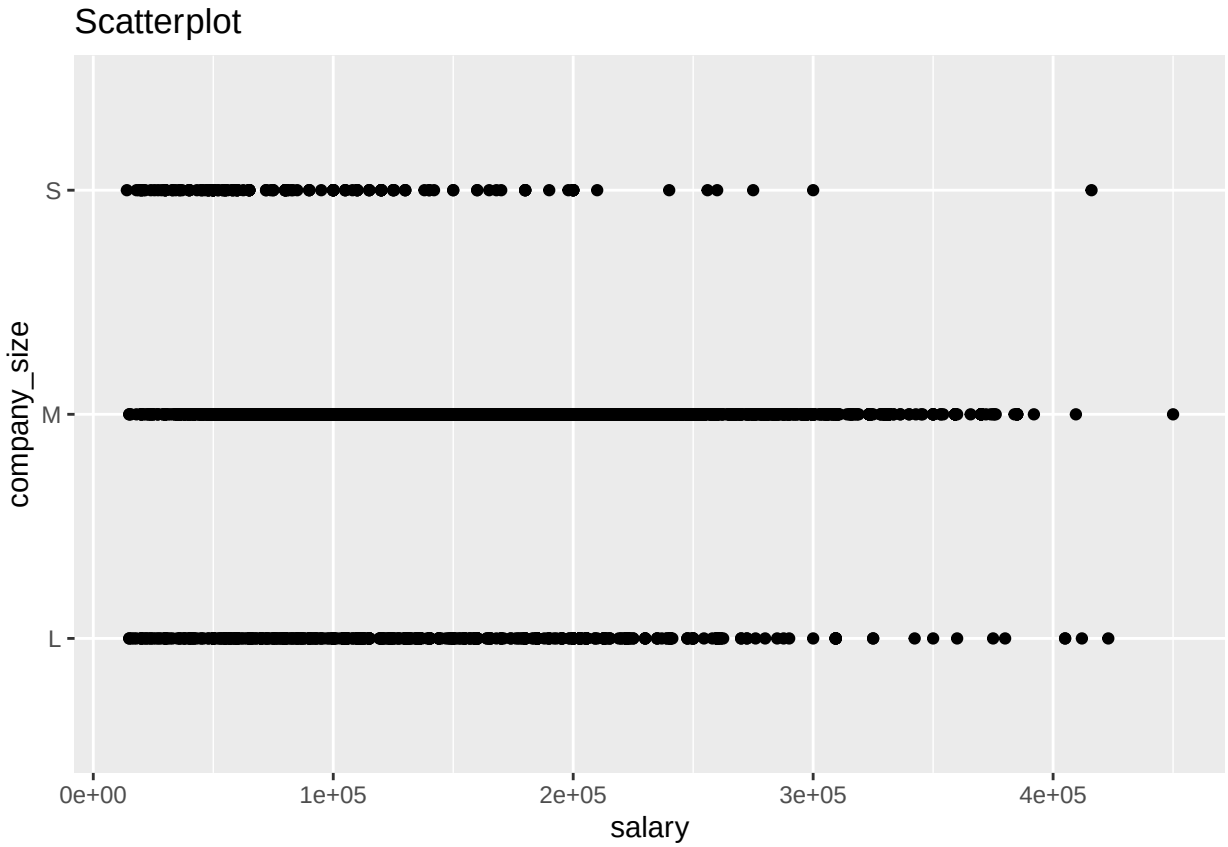
```
##   work_year            job_title                  job_category salary_currency
## 1      2023 Data DevOps Engineer              Data Engineering             EUR
## 2      2023       Data Architect Data Architecture and Modeling             USD
## 3      2023       Data Architect Data Architecture and Modeling             USD
## 4      2023       Data Scientist      Data Science and Research             USD
## 5      2023       Data Scientist      Data Science and Research             USD
## 6      2023       Data Scientist      Data Science and Research             USD
##   salary salary_in_usd employee_residence experience_level employment_type
## 1  88000         95012            Germany        Mid-level       Full-time
## 2 186000        186000      United States           Senior       Full-time
```

```
## 3  81800         81800      United States           Senior       Full-time
## 4 212000        212000      United States           Senior       Full-time
## 5  93300         93300      United States           Senior       Full-time
## 6 130000        130000      United States           Senior       Full-time
##   work_setting company_location company_size TransformedSalary
## 1       Hybrid          Germany            L          296.6479
## 2    In-person    United States            M          431.2772
## 3    In-person    United States            M          286.0070
## 4    In-person    United States            M          460.4346
## 5    In-person    United States            M          305.4505
## 6       Remote    United States            M          360.5551
```

```
library(ggplot2)
ggplot(jobs_in_data1, aes (x = salary)) +
geom_histogram (binwidth = 5, fill ="blue",color = "black") +
labs (title = "Histogram of Quantitative Variable", x = "Quantitative Variable", y = "Frequency")
```




```
# Plot a scatterplot
library(ggplot2)
ggplot (jobs_in_data1, aes (x = salary, y =company_size))+
geom_point() +
labs(title= "Scatterplot", X = "salary", Y = "company_size")
```


Scatterplot
company_size
S
M
L
0e+00
1e+05
2e+05
3e+05
4e+05
salary