

RESEARCH METHODS IN FINANCE

Quantile regression analysis of wage determinants in Vietnam

Presented by Tran Viet Hang & Tran Thanh Dat
Instructor: Dr. Nguyen Phuong Anh





Table of Contents

01 Introduction

02 Literature review

03 Methodology

04 Application
Quantile Regression in R

05 Conclusion

1. INTRODUCTION



Salary is influenced by various factors, resulting in different outcomes and wage disparity in the community.

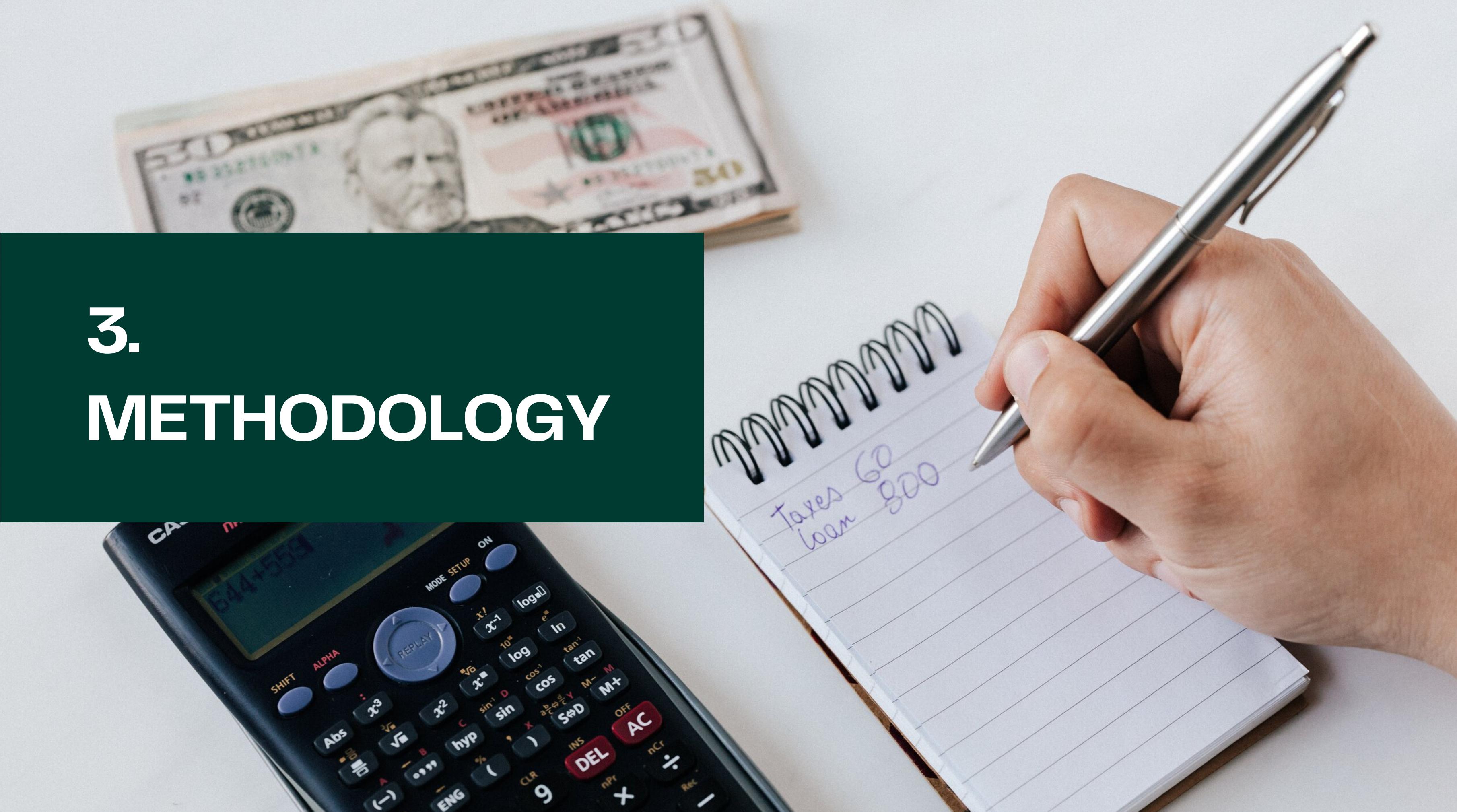
2. LITERATURE REVIEW

Mincer earnings function:

$$\ln(\omega) = \ln(\omega_0) + \rho s + \beta_1 x + \beta_2 x^2$$

- $\ln(\omega) = \ln(\omega_0) + \rho s + \beta_1 x + \beta_2 x^2$
- ω : *earnings*
- ω_0 : *earnings of someone with no education and no experience*
- s : *years of schooling*
- x : *years of potential labour market experience*
- ρ, β_1, β_2 : *regression coefficient*

3. METHODOLOGY



3. METHODOLOGY

Standard linear regression model: misleading estimates from violation of assumptions

Standard non-linear regression model: how to choose square or cubic or other polynomial terms in the equation

Quantile regressions: a more natural and flexible way to capture the relationship

3. METHODOLOGY

Linear regression model

- Model equation:

$$Y = \beta X, \quad i = 1, \dots, n$$

- Mean squared error for linear regression

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \beta x_i)^2$$

- Minimization problem:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \frac{1}{n} \sum_{i=1}^n (y_i - \beta x_i)^2$$

3. METHODOLOGY

Quantile regression model

- Model equation:

$$Y = \beta X, \quad i = 1, \dots, n$$

- Mean absolute deviation for linear regression

$$MAD = \sum_{i:y_i > \beta x_i}^n \tau |y_i - \beta x_i| + \sum_{i:y_i < \beta x_i}^n (1 - \tau) |y_i - \beta x_i|$$

- Minimization problem:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \left(\sum_{i:y_i > \beta x_i}^n \tau |y_i - \beta x_i| + \sum_{i:y_i < \beta x_i}^n (1 - \tau) |y_i - \beta x_i| \right)$$

3. METHODOLOGY

Quantile regression model

- Model equation:

$$Y = \beta X, \quad i = 1, \dots, n$$

- Mean absolute deviation for linear regression

$$MAD = \sum_{i:y_i > \beta x_i}^n 0.5|y_i - \beta x_i| + \sum_{i:y_i < \beta x_i}^n 0.5|y_i - \beta x_i|$$

- Minimization problem:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \left(\sum_{i:y_i > \beta x_i}^n \tau|y_i - \beta x_i| + \sum_{i:y_i < \beta x_i}^n (1 - \tau)|y_i - \beta x_i| \right)$$

3. METHODOLOGY

Quantile regression model

- Model equation:

$$Y = \beta X, \quad i = 1, \dots, n$$

- Mean absolute deviation for linear regression

$$MAD = \sum_{i:y_i > \beta x_i}^n 0.1|y_i - \beta x_i| + \sum_{i:y_i < \beta x_i}^n 0.9|y_i - \beta x_i|$$

- Minimization problem:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \left(\sum_{i:y_i > \beta x_i}^n \tau|y_i - \beta x_i| + \sum_{i:y_i < \beta x_i}^n (1 - \tau)|y_i - \beta x_i| \right)$$

4. Quantile Regression In R

4.1. Data overview

The screenshot shows a web browser displaying the CareerBuilder.vn website. The search bar at the top includes filters for 'Bất động...' (Real estate), 'Chứng k...' (證書), 'Ngân hà...' (Banking), and 'Hồ Chí Minh'. Below the search bar, it says '1,967 việc làm' (1,967 jobs). The first result is for 'Trưởng Phòng Kinh Doanh - Thu Nhập Từ 50 - 100 Triệu' at 'Công ty TNHH SHE GROUP' in Hồ Chí Minh, with a salary range of '\$ Lương: 50 Tr - 100 Tr VND'. The second result is for 'Business Development Manager (factory construction, industrial EPC)' at 'CareerBuilder's client' in Hồ Chí Minh, with a salary range of '\$ Lương: 50 Tr - 85 Tr VND'. The third result is for 'Giám Đốc Pháp Lý Tập Đoàn Bất Động Sản' at 'TẬP ĐOÀN NAMGROUP' in Hồ Chí Minh, with a salary range of '\$ Lương: 50 Tr - 80 Tr VND'. The fourth result is for 'Giám Đốc Kinh Doanh' at 'Công ty CP BDS LINKGROUP' in Hồ Chí Minh, with a salary range of '\$ Lương: 50 Tr - 80 Tr VND'.

Trưởng Phòng Kinh Doanh - Thu Nhập Từ 50 - 100 Triệu
Công ty TNHH SHE GROUP
\$ Lương: 50 Tr - 100 Tr VND
Hồ Chí Minh
Xem video
Lưu việc làm
14/11/2021

Business Development Manager (factory construction, industrial EPC)
CareerBuilder's client
\$ Lương: 50 Tr - 85 Tr VND
Hồ Chí Minh
Lưu việc làm
11/12/2021

Giám Đốc Pháp Lý Tập Đoàn Bất Động Sản
TẬP ĐOÀN NAMGROUP
\$ Lương: 50 Tr - 80 Tr VND
Hồ Chí Minh
Lưu việc làm
10/12/2021

Giám Đốc Kinh Doanh
Công ty CP BDS LINKGROUP
\$ Lương: 50 Tr - 80 Tr VND
Hồ Chí Minh
Lưu việc làm

4. Quantile Regression In R

4.1. Data overview

The screenshot shows a job listing on the CAREERBUILDER website. The header includes the CAREERBUILDER logo, navigation links like 'Tim Việc Làm', 'CV Hay', 'VietnamSalary', 'CareerMap', 'Cẩm Nang', 'Tính Lương', and user options like 'Đăng nhập' and 'Đăng ký'. A search bar at the top allows filtering by 'Chức danh, Kỹ năng, Tên công ty', 'Tất cả ngành nghề', and 'Tất cả địa điểm'. The main content features a large banner for 'SHE GROUP' with the tagline 'ĐỒNG HÀNH - HỢP TÁC - THÀNH CÔNG' and 'Trưởng Phòng Kinh Doanh - Thu Nhập Từ 50 - 100 Triệu'. Below the banner, there's a green button labeled 'NỘP ĐƠN ỨNG TUYỂN'. The job details section includes fields for 'Địa điểm' (Hồ Chí Minh), 'Ngày cập nhật' (14/11/2021), 'Lương' (50 Tr - 100 Tr VND), 'Ngành nghề' (Bảo hiểm, Tài chính / Đầu tư, Bất động sản), 'Kinh nghiệm' (2 - 3 Năm), 'Cấp bậc' (Quản lý), and 'Hết hạn nộp' (13/12/2021). A sidebar on the right lists other similar job opportunities.

ĐỒNG HÀNH - HỢP TÁC - THÀNH CÔNG
SHE GROUP

Trưởng Phòng Kinh Doanh - Thu Nhập Từ 50 - 100 Triệu
Công ty TNHH SHE GROUP

NỘP ĐƠN ỨNG TUYỂN

Chi tiết Tổng quan công ty

Địa điểm: Hồ Chí Minh

Ngày cập nhật: 14/11/2021

Ngành nghề: Bảo hiểm, Tài chính / Đầu tư, Bất động sản

Hình thức: Nhân viên chính thức

Lương: 50 Tr - 100 Tr VND

Kinh nghiệm: 2 - 3 Năm

Cấp bậc: Quản lý

Hết hạn nộp: 13/12/2021

PHÚC LỢI

Laptop, Phụ cấp

Chế độ bảo hiểm, Xe đưa đón

Du Lịch, Du lịch nước ngoài

CÁC CÔNG VIỆC TƯƠNG TỰ

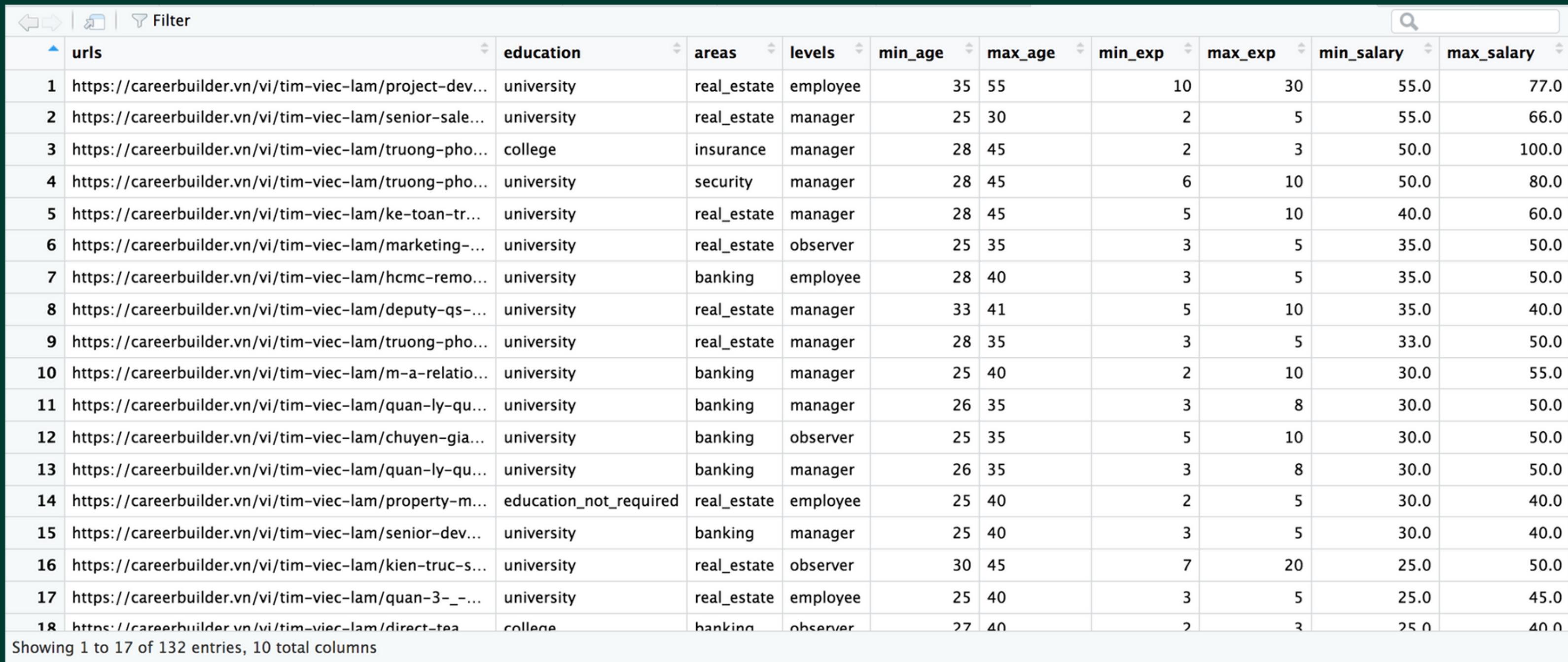
Trưởng Phòng...
CÔNG TY CỔ PHẦ...
\$ Lương:Cạnh tranh
Hồ Chí Minh

[HLI] Trưởng...
CÔNG TY CỔ PHẦ...
\$ Lương:Cạnh tranh
Hồ Chí Minh

Chuyên Viên Tư...
Tập đoàn Bất Động...
\$ Lương: 20 Tr - 100
Tr VND
Hồ Chí Minh

4. Quantile Regression In R

4.1. Data overview



The screenshot shows a data viewer interface with a toolbar at the top containing icons for back, forward, refresh, and filter, along with a search bar. The main area displays a table with 17 rows and 13 columns. The columns are labeled: urls, education, areas, levels, min_age, max_age, min_exp, max_exp, min_salary, and max_salary. The rows are numbered 1 to 18. The data includes various URLs from careerbuilder.vn, categories like university, college, and banking, and specific roles like employee, manager, and observer. Salary ranges are also listed.

	urls	education	areas	levels	min_age	max_age	min_exp	max_exp	min_salary	max_salary
1	https://careerbuilder.vn/vi/tim-viec-lam/project-dev...	university	real_estate	employee	35	55	10	30	55.0	77.0
2	https://careerbuilder.vn/vi/tim-viec-lam/senior-sale...	university	real_estate	manager	25	30	2	5	55.0	66.0
3	https://careerbuilder.vn/vi/tim-viec-lam/truong-pho...	college	insurance	manager	28	45	2	3	50.0	100.0
4	https://careerbuilder.vn/vi/tim-viec-lam/truong-pho...	university	security	manager	28	45	6	10	50.0	80.0
5	https://careerbuilder.vn/vi/tim-viec-lam/ke-toan-tr...	university	real_estate	manager	28	45	5	10	40.0	60.0
6	https://careerbuilder.vn/vi/tim-viec-lam/marketing-...	university	real_estate	observer	25	35	3	5	35.0	50.0
7	https://careerbuilder.vn/vi/tim-viec-lam/hcmc-remo...	university	banking	employee	28	40	3	5	35.0	50.0
8	https://careerbuilder.vn/vi/tim-viec-lam/deputy-qs-...	university	real_estate	manager	33	41	5	10	35.0	40.0
9	https://careerbuilder.vn/vi/tim-viec-lam/truong-pho...	university	real_estate	manager	28	35	3	5	33.0	50.0
10	https://careerbuilder.vn/vi/tim-viec-lam/m-a-relatio...	university	banking	manager	25	40	2	10	30.0	55.0
11	https://careerbuilder.vn/vi/tim-viec-lam/quan-ly-qu...	university	banking	manager	26	35	3	8	30.0	50.0
12	https://careerbuilder.vn/vi/tim-viec-lam/chuyen-gia...	university	banking	observer	25	35	5	10	30.0	50.0
13	https://careerbuilder.vn/vi/tim-viec-lam/quan-ly-qu...	university	banking	manager	26	35	3	8	30.0	50.0
14	https://careerbuilder.vn/vi/tim-viec-lam/property-m...	education_not_required	real_estate	employee	25	40	2	5	30.0	40.0
15	https://careerbuilder.vn/vi/tim-viec-lam/senior-dev...	university	banking	manager	25	40	3	5	30.0	40.0
16	https://careerbuilder.vn/vi/tim-viec-lam/kien-truc-s...	university	real_estate	observer	30	45	7	20	25.0	50.0
17	https://careerbuilder.vn/vi/tim-viec-lam/quan-3-_-...	university	real_estate	employee	25	40	3	5	25.0	45.0
18	https://careerbuilder.vn/vi/tim-viec-lam/direct-tea...	college	banking	observer	27	40	2	3	25.0	40.0

Showing 1 to 17 of 132 entries, 10 total columns

urls	male	female	university	college	intermediate	education_not_required	banking	insurance	security	real_estate	employee	observer	manager	chief_manager	age	min_age	min_exp	min_salary	education	areas	levels	gender
1 https://careerbuilder.vn/vi/tim-viec-lam/project...	0	0	1	0	0	0	0	0	0	1	1	0	0	0	35 - 55	35	10	55.0	university	real_estate	employee	0
2 https://careerbuilder.vn/vi/tim-viec-lam/senior...	0	0	1	0	0	0	0	0	0	1	0	0	1	0	25 - 30	25	2	55.0	university	real_estate	manager	0
3 https://careerbuilder.vn/vi/tim-viec-lam/truong...	0	0	0	1	0	0	0	0	1	0	1	0	1	0	28 - 45	28	2	50.0	college	insurance	manager	0
4 https://careerbuilder.vn/vi/tim-viec-lam/truong...	0	0	1	0	0	0	0	0	1	1	0	0	1	0	28 - 45	28	6	50.0	university	security	manager	0
5 https://careerbuilder.vn/vi/tim-viec-lam/ke-toa...	0	0	1	0	0	0	0	0	0	1	0	0	1	0	28 - 45	28	5	40.0	university	real_estate	manager	0
6 https://careerbuilder.vn/vi/tim-viec-lam/market...	0	0	1	0	0	0	0	0	0	1	0	1	0	0	25 - 35	25	3	35.0	university	real_estate	observer	0
7 https://careerbuilder.vn/vi/tim-viec-lam/hcmc-...	0	0	1	0	0	0	1	0	0	0	1	0	0	0	28 - 40	28	3	35.0	university	banking	employee	0
8 https://careerbuilder.vn/vi/tim-viec-lam/deputy...	0	0	1	0	0	0	0	0	0	1	0	0	1	0	33 - 41	33	5	35.0	university	real_estate	manager	0
9 https://careerbuilder.vn/vi/tim-viec-lam/truong...	0	0	1	0	0	0	0	0	0	1	0	0	1	0	28 - 35	28	3	33.0	university	real_estate	manager	0
10 https://careerbuilder.vn/vi/tim-viec-lam/m-a-r...	0	0	1	0	0	0	1	0	0	0	0	0	1	0	25 - 40	25	2	30.0	university	banking	manager	0
11 https://careerbuilder.vn/vi/tim-viec-lam/quan-l...	0	0	1	0	0	0	1	1	0	0	0	0	1	0	26 - 35	26	3	30.0	university	banking	manager	0
12 https://careerbuilder.vn/vi/tim-viec-lam/chuyen...	0	0	1	0	0	0	1	0	0	0	0	0	1	0	25 - 35	25	5	30.0	university	banking	observer	0
13 https://careerbuilder.vn/vi/tim-viec-lam/quan-l...	0	0	1	0	0	0	1	1	0	0	0	0	1	0	26 - 35	26	3	30.0	university	banking	manager	0
14 https://careerbuilder.vn/vi/tim-viec-lam/proper...	0	0	0	0	0	1	0	0	0	1	1	0	0	0	25 - 40	25	2	30.0	education_not_required	real_estate	employee	0
15 https://careerbuilder.vn/vi/tim-viec-lam/senior...	0	0	1	0	0	0	1	0	0	0	0	0	1	0	25 - 40	25	3	30.0	university	banking	manager	0
16 https://careerbuilder.vn/vi/tim-viec-lam/kien-tr...	0	0	1	0	0	0	0	0	0	1	0	0	1	0	30 - 45	30	7	25.0	university	real_estate	observer	0
17 https://careerbuilder.vn/vi/tim-viec-lam/quan-...	0	0	1	0	0	0	0	0	0	1	1	0	0	0	25 - 40	25	3	25.0	university	real_estate	employee	0
18 https://careerbuilder.vn/vi/tim-viec-lam/direct-...	0	0	0	1	0	0	1	1	0	0	0	0	1	0	27 - 40	27	2	25.0	college	banking	observer	0
19 https://careerbuilder.vn/vi/tim-viec-lam/ke-toa...	0	0	1	0	0	0	0	0	0	1	0	0	0	1	31 - 37	31	5	25.0	university	real_estate	manager	0
20 https://careerbuilder.vn/vi/tim-viec-lam/truong...	0	0	1	0	0	0	0	0	0	1	0	0	1	0	35 - 50	35	0	25.0	university	real_estate	manager	0
21 https://careerbuilder.vn/vi/tim-viec-lam/chuyen...	0	0	1	0	0	0	0	0	0	1	1	0	0	0	27 - 45	27	3	23.0	university	real_estate	employee	0
22 https://careerbuilder.vn/vi/tim-viec-lam/sales-t...	0	0	0	0	0	1	0	1	1	0	0	0	1	0	22 - 40	22	1	22.0	education_not_required	insurance	observer	0
23 https://careerbuilder.vn/vi/tim-viec-lam/sales-...	0	0	0	1	0	0	0	1	0	1	0	0	0	1	25 - 40	25	1	22.0	college	banking	manager	0
24 https://careerbuilder.vn/vi/tim-viec-lam/chuyen...	0	0	1	0	0	0	0	0	0	1	1	0	0	0	30 - 40	30	3	22.0	university	real_estate	employee	0
25 https://careerbuilder.vn/vi/tim-viec-lam/fx-bus...	0	0	0	1	0	0	0	1	0	0	0	0	0	0	25 - 55	25	3	22.0	college	banking	observer	0
26 https://careerbuilder.vn/vi/tim-viec-lam/truong...	0	0	0	1	0	0	0	0	0	1	0	0	1	0	26 - 38	26	1	20.0	college	real_estate	observer	0
27 https://careerbuilder.vn/vi/tim-viec-lam/truong...	0	0	0	0	1	0	0	0	1	0	0	1	0	0	26 - 40	26	1	20.0	intermediate	insurance	manager	0
28 https://careerbuilder.vn/vi/tim-viec-lam/chuyen...	0	0	1	0	0	0	0	0	1	0	1	0	0	0	20 - 30	20	2	20.0	university	security	employee	0
29 https://careerbuilder.vn/vi/tim-viec-lam/custo...	0	0	1	0	0	0	0	1	0	0	0	1	0	0	25 - 45	25	3	20.0	university	banking	employee	0
30 https://careerbuilder.vn/vi/tim-viec-lam/quan-...	0	0	1	0	0	0	0	1	1	0	0	1	0	0	29 - 37	29	4	20.0	university	banking	employee	0
31 https://careerbuilder.vn/vi/tim-viec-lam/chuyen...	0	0	0	1	0	0	0	0	0	1	1	0	0	0	25 - 35	25	0	20.0	college	real_estate	employee	0
32 https://careerbuilder.vn/vi/tim-viec-lam/chuyen...	0	0	1	0	0	0	0	0	1	1	0	0	0	0	25 - 35	25	3	20.0	university	real_estate	employee	0
33 https://careerbuilder.vn/vi/tim-viec-lam/truong...	0	0	1	0	0	0	0	0	1	0	0	0	0	1	30 - 40	30	3	20.0	university	security	manager	0
34 https://careerbuilder.vn/vi/tim-viec-lam/hcm-c...	0	0	1	0	0	0	0	1	0	1	0	1	0	0	23 - 35	23	2	20.0	university	banking	employee	0
35 https://careerbuilder.vn/vi/tim-viec-lam/senior...	0	0	1	0	0	0	0	0	0	1	1	0	0	0	25 - 35	25	0	20.0	university	real_estate	employee	0
36 https://careerbuilder.vn/vi/tim-viec-lam/ke-toa...	0	0	1	0	0	0	0	0	0	1	0	0	0	1	28 - 35	28	3	20.0	university	real_estate	manager	0
37 https://careerbuilder.vn/vi/tim-viec-lam/mall-l...	0	0	1	0	0	0	0	0	0	1	0	0	1	0	26 - 40	26	3	20.0	university	real_estate	observer	0
38 https://careerbuilder.vn/vi/tim-viec-lam/kiem-t...	0	0	1	0	0	0	0	0	1	0	0	1	0	0	30 - 38	30	5	20.0	university	real_estate	observer	0
39 https://careerbuilder.vn/vi/tim-viec-lam/tro-ly...	0	0	1	0	0	0	0	0	1	0	0	0	1	0	28 - 35	28	3	20.0	university	real_estate	manager	0
40 https://careerbuilder.vn/vi/tim-viec-lam/giam-...	0	0	1	0	0	0	0	0	0	1	0	0	0	1	27 - 35	27	2	18.0	university	real_estate	manager	0



4.2.

DATA EXPLORATORY

4. Quantile Regression In R

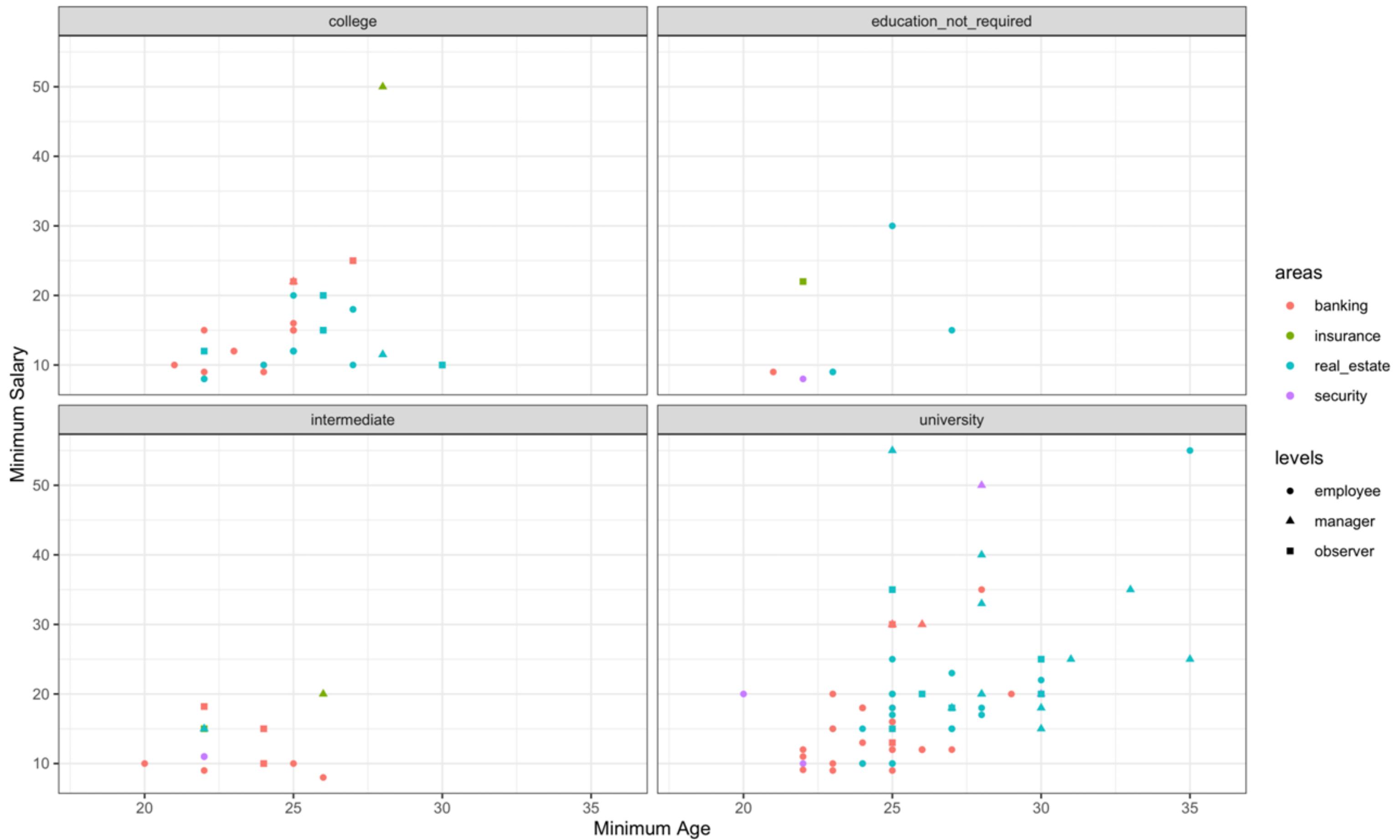
4.2. Data exploratory

```
1 library(tidyverse)
2 ## Salary ~ Age
3 ggplot(df,
4         aes(x=as.numeric(min_age),
5               y=min_salary,
6               color=areas,
7               shape=levels)) +
8     geom_point() +
9     facet_wrap(~education) +
10    labs(title="CareerBuilder Ho Chi Minh City Jobs in 2021",
11          x="Minimum Age",
12          y="Minimum Salary") +
13    theme_bw() + scale_x_continuous(limits=c(18,36))
```

QUANTILE REGRESSION

WAGE DETERMINANTS

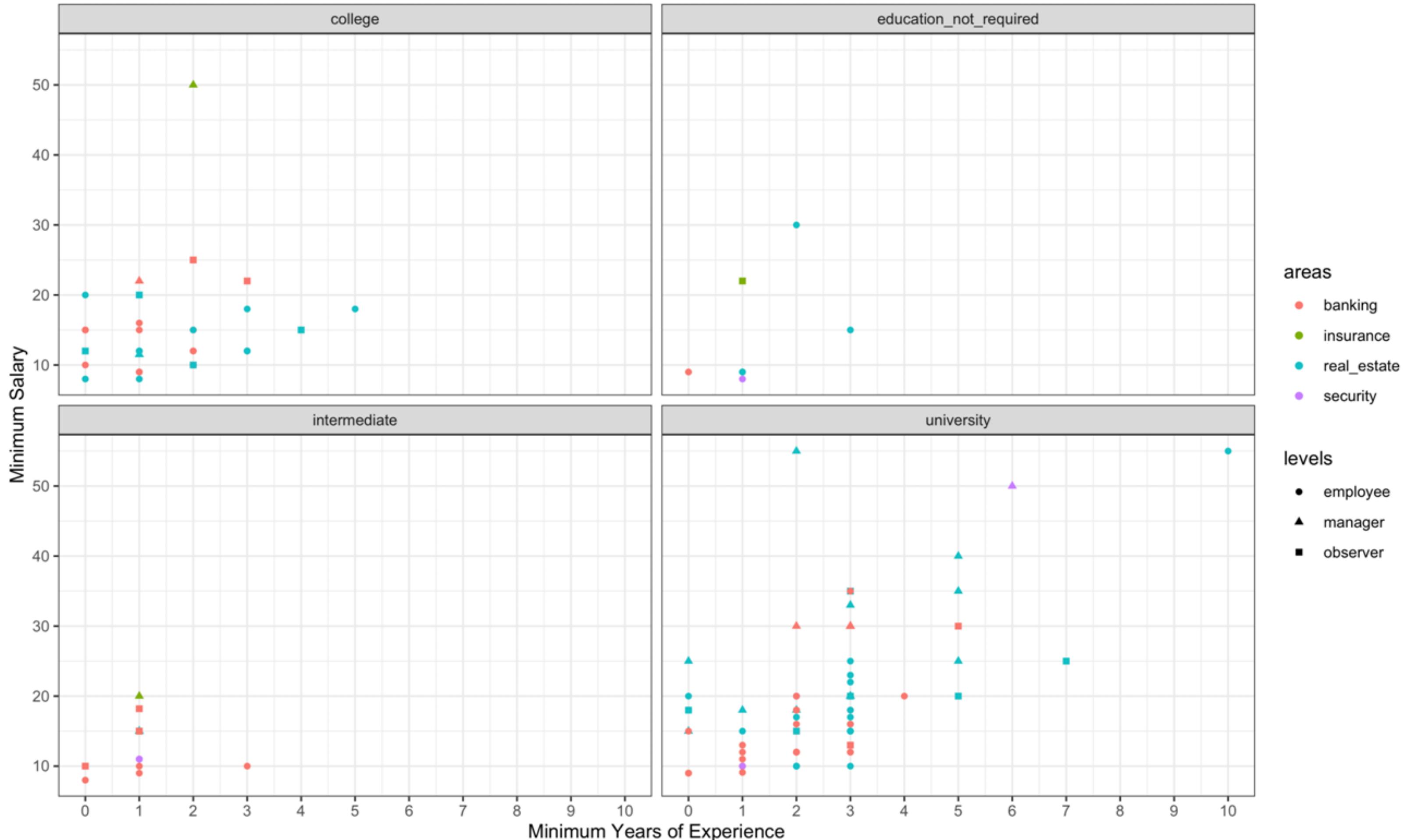
CareerBuilder Ho Chi Minh City Jobs in 2021



QUANTILE REGRESSION

WAGE DETERMINANTS

CareerBuilder Ho Chi Minh City Jobs in 2021





4.3. DATA MODELLING

DATA MODELLING

With Linear Regression Model

Call:

```
lm(formula = min_salary ~ university + college + intermediate +  
  education_not_required + banking + insurance + security +  
  real_estate + employee + observer + manager + min_age + min_exp,  
  data = df)
```

Residuals:

Min	1Q	Median	3Q	Max
-9.3386	-3.8286	-0.6764	2.4357	29.9892

Coefficients: (2 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	13.1401	8.6069	1.527	0.12947
university	-0.2435	2.6976	-0.090	0.92823
college	-2.1271	2.7371	-0.777	0.43861
intermediate	-3.3352	2.9914	-1.115	0.26712
education_not_required	NA	NA	NA	NA
banking	3.0140	3.1842	0.947	0.34576
insurance	5.4813	2.0157	2.719	0.00751 **
security	3.1239	2.4575	1.271	0.20613
real_estate	3.7259	3.4863	1.069	0.28733
employee	-9.5280	1.7146	-5.557	1.68e-07 ***
observer	-6.4152	2.0629	-3.110	0.00234 **
manager	NA	NA	NA	NA
min_age	0.1229	0.3052	0.403	0.68780
min_exp	2.6574	0.4484	5.927	3.04e-08 ***

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
	.	1		

Residual standard error: 6.349 on 120 degrees of freedom

Multiple R-squared: 0.5527, Adjusted R-squared: 0.5117

F-statistic: 13.48 on 11 and 120 DF, p-value: < 2.2e-16

Call:

```
lm(formula = min_salary ~ areas + 0, data = df)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-11.750	-5.889	-3.231	2.383	36.769

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
areasbanking	15.617	1.218	12.826	< 2e-16 ***
areasinsurance	26.750	4.474	5.979	2.09e-08 ***
areasreal_estate	18.231	1.110	16.427	< 2e-16 ***
areassecurity	16.889	2.982	5.663	9.34e-08 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.947 on 128 degrees of freedom

Multiple R-squared: 0.7969, Adjusted R-squared: 0.7905

F-statistic: 125.5 on 4 and 128 DF, p-value: < 2.2e-16

The average salary of people working in the insurance field stands at the highest with 26 million VND per month. Other fields seem to be consistent in the range from 15 to 18 million per month.

Managers can earn more than 26 million VND per month on average

Call:

```
lm(formula = min_salary ~ levels + 0, data = df)
```

Residuals:

Min	1Q	Median	3Q	Max
-15.125	-4.473	-2.049	3.375	40.527

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
levelemployee	14.4730	0.8329	17.38	<2e-16 ***	
levelsmanager	26.6250	1.6039	16.60	<2e-16 ***	
levelsobserver	18.9579	1.8027	10.52	<2e-16 ***	

Signif. codes:	0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				

Residual standard error: 7.858 on 129 degrees of freedom

Multiple R-squared: 0.8421, Adjusted R-squared: 0.8385

F-statistic: 229.4 on 3 and 129 DF, p-value: < 2.2e-16

Multiple R-squared is only ~30%

Call:

```
lm(formula = min_salary ~ min_exp, data = df)
```

Residuals:

Min	1Q	Median	3Q	Max
-10.520	-5.010	-1.937	2.772	37.772

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	10.6453	1.0983	9.693	< 2e-16 ***
min_exp	3.2915	0.4323	7.614	4.81e-12 ***

Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’
	0.1 ‘ ’	1		

Residual standard error: 7.585 on 130 degrees of freedom

Multiple R-squared: 0.3084, Adjusted R-squared: 0.3031

F-statistic: 57.97 on 1 and 130 DF, p-value: 4.812e-12

DATA MODELLING

With Quantile Regression Model

```
1 library(quantreg)
2 quantreg.all <- rq(min_salary ~ banking+
3                      insurance+
4                      security+
5                      real_estate+
6                      employee+
7                      manager+
8                      min_age+
9                      min_exp,
10                     tau = 0.5, data=df)
11 quantreg <- summary(quantreg.all)
```

DATA MODELLING

With Quantile Regression Model

```
Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +  
employee + manager + min_age + min_exp, tau = 0.5, data = df)
```

```
tau: [1] 0.5
```

Coefficients:

	coefficients	lower bd	upper bd
(Intercept)	5.05000	-10.45416	13.63593
banking	2.45000	1.24067	5.80138
insurance	6.00000	1.34777	7.03562
security	2.86667	0.33738	6.54666
real_estate	1.45000	-1.08622	5.95579
employee	-4.75000	-5.77669	-0.84132
manager	1.66667	-0.25915	9.75462
min_age	0.25000	-0.19854	0.90521
min_exp	2.33333	0.79159	3.42946

DATA MODELLING

With Quantile Regression Model

```
1 library(quantreg)
2 quantreg.all <- rq(min_salary~banking+
3                     insurance+
4                     real_estate+
5                     security+
6                     employee+
7                     manager+
8                     min_exp, tau = seq(0.05, 0.95, by = 0.05), data=df)
9 quantreg <- summary(quantreg.all)
```

```

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.1

Coefficients:
 coefficients lower bd upper bd
(Intercept) 6.900000e+00 4.143990e+00 1.797693e+008
banking     1.000000e-01 -1.797693e+008 1.797693e+008
insurance   5.000000e+00 -1.797693e+008 5.786620e+000
security    1.100000e+00 -1.797693e+008 2.445500e+000
real_estate 1.000000e-01 -1.797693e+008 1.940900e+000
employee   -1.000000e+00 -3.075280e+000 1.797693e+008
manager     6.000000e+00 7.788400e-01 7.174680e+000
min_exp     2.000000e+00 -5.528200e-01 3.064910e+000

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.6

Coefficients:
 coefficients lower bd upper bd
(Intercept) 9.80000 5.30441 14.10412
banking      3.20000 1.51740 5.61195
insurance    4.20000 1.85751 17.09395
security     2.20000 -0.48931 4.92328
real_estate  3.20000 0.75466 5.43400
employee    -4.00000 -7.49851 -2.04760
manager      3.80000 1.40180 11.68450
min_exp      3.00000 1.32503 3.90665

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.2

Coefficients:
 coefficients lower bd upper bd
(Intercept) 7.333330e+00 3.078880e+00 1.082800e+01
banking     2.000000e+00 -5.692800e-01 4.559040e+00
insurance   4.500000e+00 2.820900e-01 6.926100e+00
security    1.000000e+00 -1.797693e+008 3.611690e+000
real_estate 8.333300e-01 -2.479110e+000 3.126570e+000
employee   -2.500000e+00 -4.207470e+000 -1.309500e-01
manager     5.166670e+00 1.815760e+000 8.331630e+000
min_exp     2.166670e+00 1.340580e+000 2.989820e+000

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.7

Coefficients:
 coefficients lower bd upper bd
(Intercept) 9.80000 6.41724 16.63576
banking      3.20000 0.27825 7.35328
insurance    6.00000 -0.25857 6.76536
security     2.20000 0.04554 9.39610
real_estate  3.20000 0.81090 9.75759
employee    -4.00000 -7.57586 -2.46534
manager      8.00000 2.46203 12.24209
min_exp      3.00000 1.34493 4.06416

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.3

Coefficients:
 coefficients lower bd upper bd
(Intercept) 7.00000 3.85431 12.89569
banking     3.00000 -1.16239 5.81377
insurance   4.33333 1.57825 6.50362
security    2.16667 -0.84060 4.67576
real_estate 2.16667 -2.36517 4.92364
employee   -2.33333 -6.48170 0.81028
manager     4.33333 -0.45251 7.48417
min_exp     2.16667 1.06395 3.36323

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.8

Coefficients:
 coefficients lower bd upper bd
(Intercept) 8.285710e+00 3.888500e+00 1.859841e+01
banking     5.285710e+00 -2.730730e+00 8.724400e+00
insurance   6.000000e+00 4.067870e+00 2.002190e+01
security    4.428570e+00 -5.560800e-01 1.797693e+008
real_estate 8.428570e+00 -1.863280e+00 1.357911e+01
employee   -4.571430e+00 -7.757280e+00 -2.062360e+00
manager     8.285710e+00 -1.679520e+00 1.132013e+01
min_exp     3.285710e+00 1.846050e+00 4.436910e+00

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.4

Coefficients:
 coefficients lower bd upper bd
(Intercept) 7.00000 4.69588 13.06507
banking     3.00000 -0.48956 5.88956
insurance   6.00000 3.19607 7.88393
security    4.00000 -0.80820 6.14466
real_estate 4.00000 -0.66769 6.49440
employee   -2.00000 -6.44977 -0.69835
manager     4.00000 -0.30535 8.55349
min_exp     2.00000 1.60687 4.06663

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.9

Coefficients:
 coefficients lower bd upper bd
(Intercept) 1.250000e+01 -1.797693e+008 2.056307e+01
banking     0.000000e+00 -1.797693e+008 1.797693e+008
insurance   5.100000e+00 2.847440e+00 1.797693e+008
security    2.000000e+00 -2.303070e+00 1.797693e+008
real_estate 7.400000e+00 5.465950e+00 1.797693e+008
employee   -1.900000e+00 -1.797693e+008 8.332200e-01
manager     1.010000e+01 3.144220e+00 2.923894e+01
min_exp     3.700000e+00 1.924770e+00 3.939890e+00

Call: rq(formula = min_salary ~ banking + insurance + security + real_estate +
  employee + manager + min_exp, tau = seq(0.1, 0.9, 0.1), data = df)
tau: [1] 0.5

Coefficients:
 coefficients lower bd upper bd
(Intercept) 9.80000 5.03600 12.37840
banking     3.00000 1.44096 4.29920
insurance   6.40000 1.65971 6.90508
security    3.00000 -0.44550 6.50071
real_estate 2.20000 -0.09493 6.28478
employee   -4.20000 -9.68673 -0.96214
manager     3.60000 0.55698 8.65813
min_exp     2.40000 1.28363 3.99523

```

Fitted models are hard to combine, manipulate and visualize

DATA MODELING

```
15 library(broom)
16 quantreg.tidied=tidy(quantreg.all)
```

Take each model and turn it into a tidy data frame of coefficients using the tidy function in the broom package

	term	estimate	conf.low	conf.high	tau
1	(Intercept)	6.9000000	-1.797693e+308	1.797693e+308	0.1
2	banking	0.1000000	-1.797693e+308	1.797693e+308	0.1
3	insurance	5.0000000	-1.797693e+308	5.949706e+00	0.1
4	security	1.1000000	-1.797693e+308	3.568486e+00	0.1
5	real_estate	0.1000000	-1.797693e+308	1.797693e+308	0.1
6	employee	-1.0000000	-3.361972e+00	1.797693e+308	0.1
7	manager	6.0000000	5.199494e-01	1.797693e+308	0.1
8	min_exp	2.0000000	-9.313955e-01	3.115781e+00	0.1
9	(Intercept)	7.3333333	2.428025e+00	1.201996e+01	0.2
10	banking	2.0000000	-1.797693e+308	5.342788e+00	0.2
11	insurance	4.5000000	3.615817e-02	7.339254e+00	0.2
12	security	1.0000000	-1.797693e+308	4.229334e+00	0.2
13	real_estate	0.8333333	-2.785968e+00	4.705294e+00	0.2
14	employee	-2.5000000	-4.480136e+00	1.088633e+16	0.2
15	manager	5.1666667	1.067677e+00	8.745859e+00	0.2
16	min_exp	2.1666667	3.142852e-01	4.035000e+00	0.2
17	(Intercept)	7.0000000	2.185163e+00	1.410371e+01	0.3
18	banking	3.0000000	-1.438803e+00	7.239632e+00	0.3
19	insurance	4.3333333	1.363705e-01	7.213622e+00	0.3
20	security	2.1666667	-1.007568e+00	5.351189e+00	0.3
21	real_estate	2.1666667	-3.620216e+00	5.963888e+00	0.3
22	employee	-2.3333333	-6.835407e+00	1.190649e+00	0.3
23	manager	4.3333333	-1.586672e+00	9.291619e+00	0.3
24	min_exp	2.1666667	6.652680e-01	5.077960e+00	0.3
25	(Intercept)	7.0000000	4.131536e+00	1.341779e+01	0.4
26	banking	3.0000000	-8.178849e-01	7.244712e+00	0.4
27	insurance	6.0000000	2.225153e+00	9.188987e+00	0.4
28	security	4.0000000	-1.302301e+00	6.471780e+00	0.4
29	real_estate	4.0000000	-2.602678e+00	7.053416e+00	0.4
30	employee	-2.0000000	-6.722862e+00	1.176265e+00	0.4
31	manager	4.0000000	-5.500137e-01	1.017708e+01	0.4

Showing 1 to 32 of 72 entries, 5 total columns

DATA MODELING

Turn the long table into a wide data frame using the function spread

```
18 library(tidyr)
19 quantreg.wide=data_tidied %>% select(term,tau,estimate) %>% spread(tau,estimate)
```

	term	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
1	(Intercept)	4.9	6.9	7.000000e+00	7.333333	6.000000	7.000000	9	7	9.333333	9.8	9.8	9.8	9.8	9.8	11.8	8.285714	10.250	12.5	18.5
2	banking	1.1	0.1	1.000000e+00	2.000000	4.000000	3.000000	2	3	2.666667	3.0	3.2	3.2	3.2	3.2	2.2	5.285714	2.875	0.0	0.0
3	employee	0.0	-1.0	-1.000000e+00	-2.500000	-2.666667	-2.333333	-3	-2	-4.000000	-4.2	-4.0	-4.0	-4.0	-4.0	-5.0	-4.571429	-3.125	-1.9	-4.5
4	insurance	2.0	5.0	5.000000e+00	4.500000	4.666667	4.333333	5	6	5.333333	6.4	6.0	4.2	6.0	6.0	6.0	6.000000	5.125	5.1	1.0
5	manager	3.5	6.0	6.000000e+00	5.166667	4.666667	4.333333	3	4	3.000000	3.6	3.8	3.8	7.0	8.0	9.0	8.285714	8.500	10.1	17.5
6	min_exp	2.0	2.0	2.000000e+00	2.166667	2.333333	2.166667	2	2	2.333333	2.4	2.4	3.0	3.0	3.0	3.0	3.285714	3.375	3.7	3.5
7	real_estate	1.1	0.1	1.480297e-17	0.833333	2.333333	2.166667	2	4	2.666667	2.2	3.0	3.2	3.2	3.2	4.2	8.428571	7.750	7.4	6.0
8	security	1.1	1.1	1.000000e+00	1.000000	2.333333	2.166667	2	4	3.333333	3.0	2.8	2.2	2.2	2.2	1.2	4.428571	3.250	2.0	-1.0

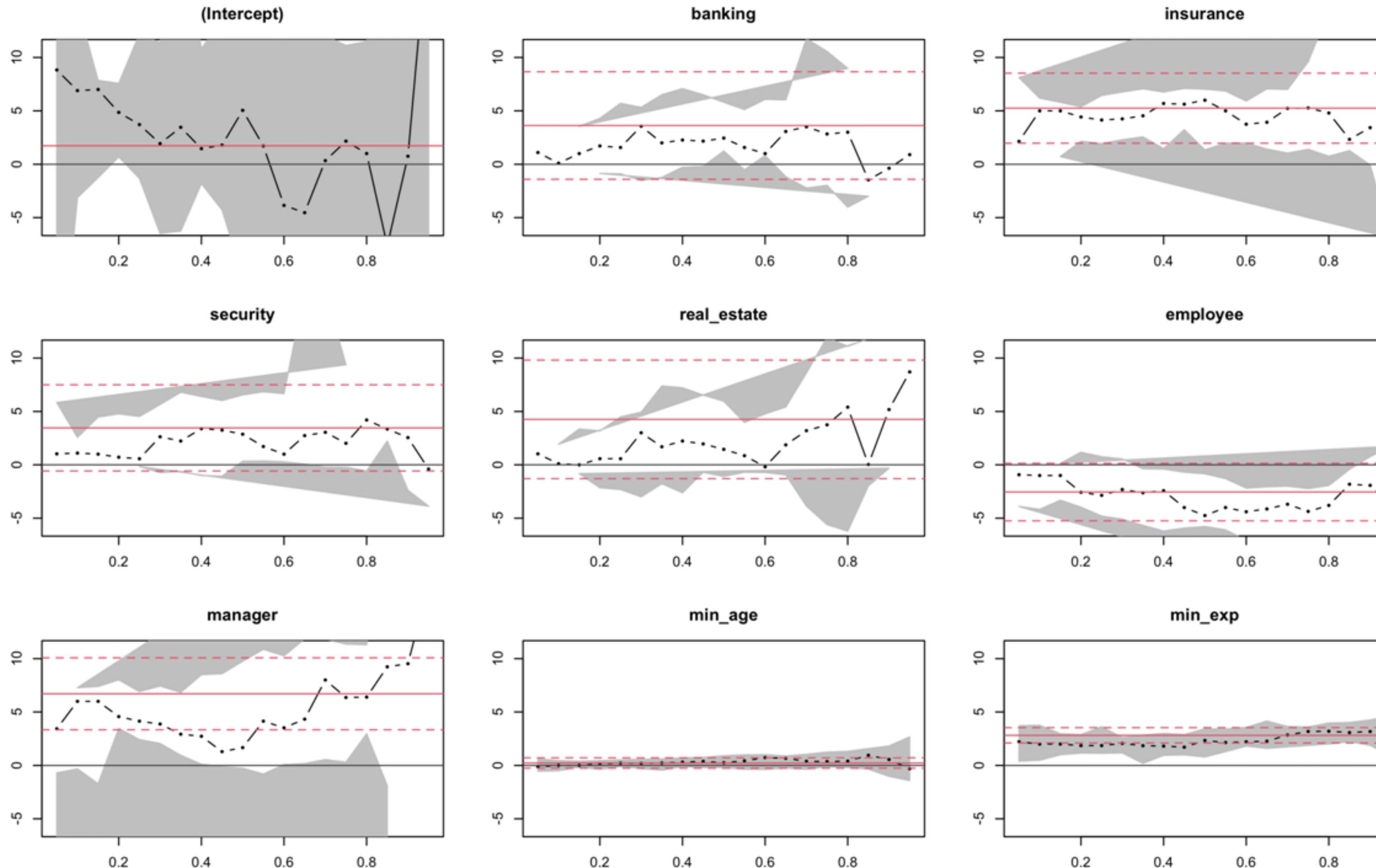
The data frame compared with the confidence interval of the linear regression

	Quantile Regression																			Linear Regression			
	term	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	estimate	std.error	confidence interval
1	(Intercept)	4.90	6.90	7.00	7.33	6.00	7.00	9.00	7.00	9.33	9.80	9.80	9.80	9.80	9.80	11.80	8.29	10.25	12.50	18.5*	6.86	3.55	(-0.24, 13.96)
2	banking	1.10	0.10	1.00	2.00	4.00	3.00	2.00	3.00	2.67	3.00	3.20	3.20	3.20	3.20	2.20	5.29	2.88	0.00	0.00	3.75*	3.05	(-2.35, 9.85)
3	insurance	2.00	5.00	5.00	4.50	4.67	4.33	5.00	6.00	5.33	6.40	6.00	4.20	6.00	6.00	6.00	6.00	5.13	5.10	1.00	5.39	1.98	(1.43, 9.35)
4	security	1.10	1.10	1.00	1.00	2.33	2.17	2.00	4.00	3.33	3.00	2.80	2.20	2.20	2.20	1.20	4.43	3.25	2.00	-1.00	3.22	2.43	(-1.64, 8.08)
5	real_estate	1.10	0.10	0.00	0.83	2.33	2.17	2.00	4.00	2.67	2.20	3.00	3.20	3.20	3.20	4.20	8.43	7.75	7.40	6.00	4.63*	3.33	(-2.03, 11.29)
6	employee	0.00	-1.00	-1.00	-2.50	-2.67	-2.33	-3.00	-2.00	-4.00	-4.20	-4.00	-4.00	-4.00	-4.00	-5.00	-4.57	-3.13	-1.90	-4.50	-2.64*	1.63	(-5.9, 0.62)
7	manager	3.50	6.00	6.00	5.17	4.67	4.33	3.00	4.00	3.00	3.60	3.80	3.80	7.00	8.00	9.00	8.29	8.50	10.10	17.5*	7.11*	1.98	(3.15, 11.07)
8	min_exp	2*	2*	2*	2.17*	2.33*	2.17*	2*	2*	2.33	2.40	2.40	3.00	3.00	3.00	3.00	3.28*	3.38*	3.7*	3.5*	2.99*	0.38	(2.23, 3.15)

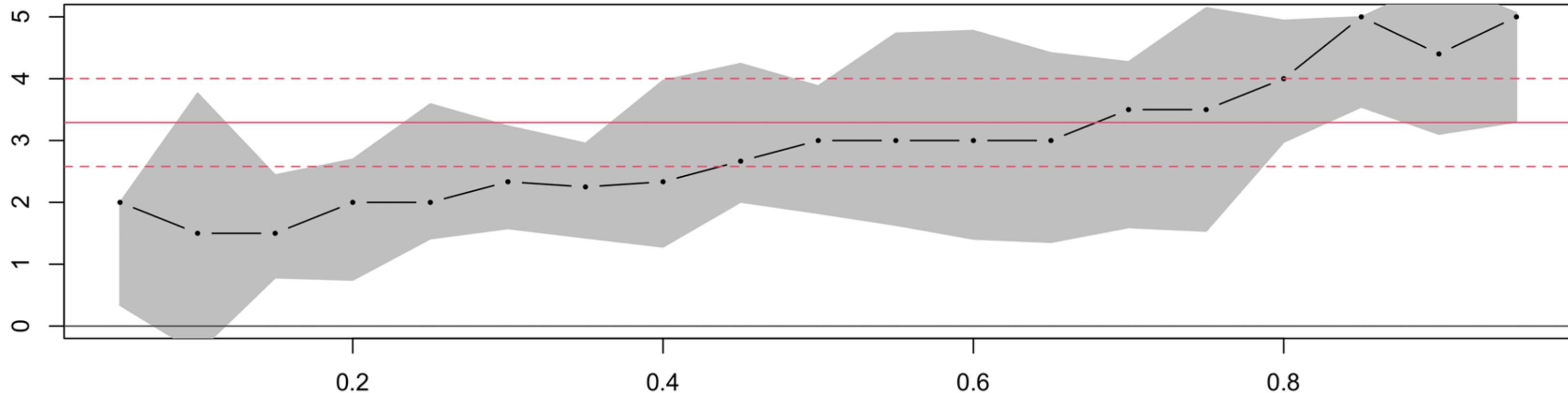
QUANTILE REGRESSION

WAGE DETERMINANTS

- Linear: banking, insurance, security, real_estate, employee, manager, min_age
- Quantile: min_exp

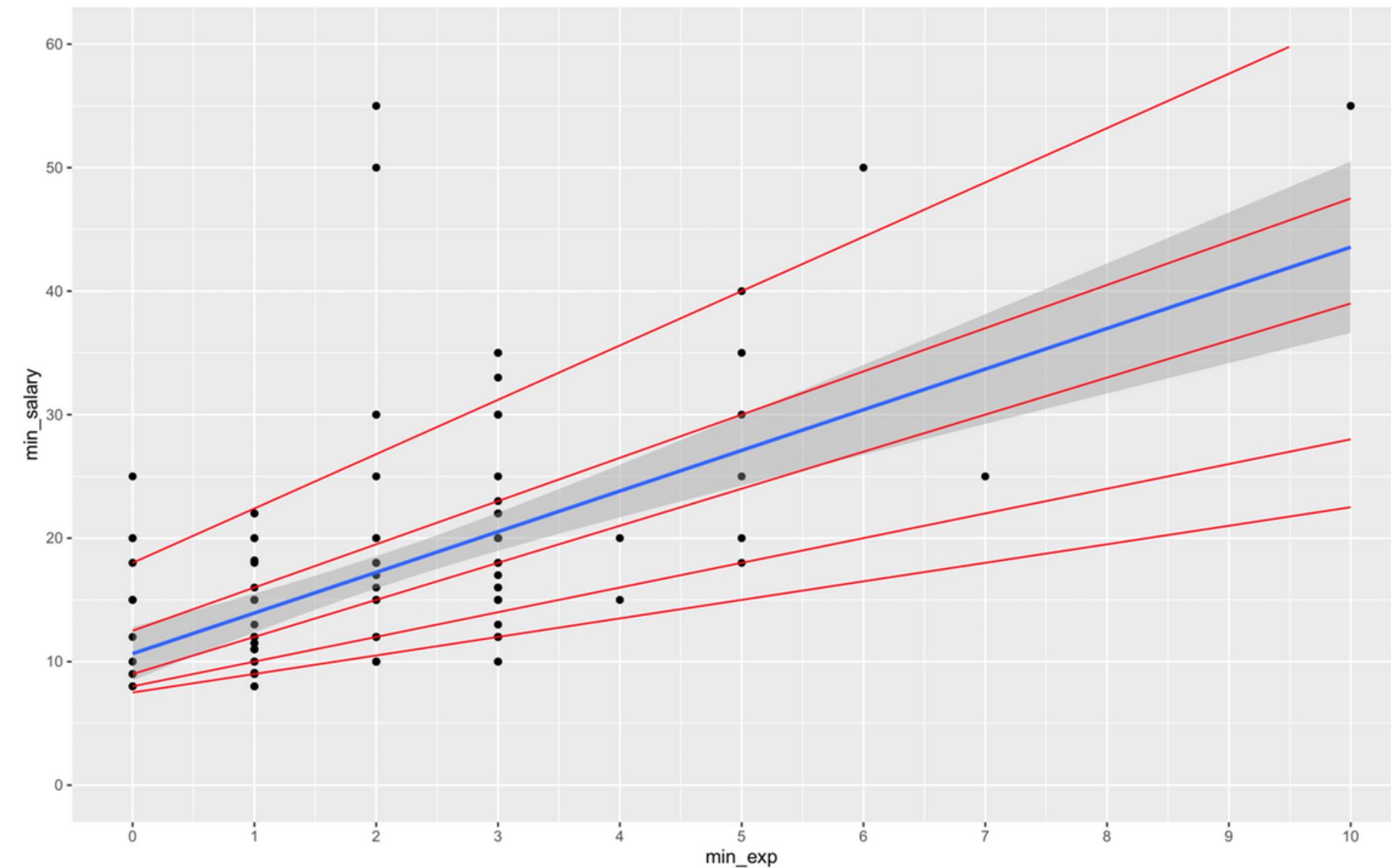


min_exp



Quantile regression is
statistically significant from 0 to
0.5 and from 0.8 to 1

5. Conclusion



Linear model: the increase of 3 million VND for 1 more year of experience.

Quantile model: the increase of salary for 1 more year of experience varies for salary less than 16 million or greater than 33 million.

Thank you for
listening!