UBS (один з найбільших швейцарських банків світу) підготував звіт, у якому порівнюються ціни, зарплати та інші економічні показники в містах по всьому світу. Дані знаходяться в файлі prices-and-earnings. Змінні, які були виміряні в 72 містах, це вартість життя, вартість харчування, середня погодинна заробітна плата та інші.

Цікаво те, що звіт UBS пропонує зручний підхід: вимірювати рівень життя по тому, скільки часу потрібно працівнику із середньою чистою зарплатою, щоб дозволити собі Біг Мак. Любителям фастфуду найкраще жити в Чикаго, Торонто та Токіо, де потрібно лише 12 хвилин на роботі, щоб дозволити собі Біг Мак. Навпаки, працівникам доводиться працювати більше двох годин, щоб поласувати гамбургерами в Мехіко, Джакарті та Найробі.

**Об’єм вибірки:** 72

**Назви змінних:**

* City: Назва міста.
* Food\_Costs: Вартість хачування
* Womens\_Clothing: Вартість жіночого одягу
* Mens\_Clothing: Вартість чоловічого одягу
* iPhone\_4S(hr): Вартість айфону
* Clothing\_index: Індекс вартості одягу
* Hours\_Worked: Cередня кількість робочих годин
* Wage\_Gross: Середня зарплатня до утримання податків
* Wage\_Net: Середня зарплатня після утримання податків
* Vacation\_Days: Середня кількість днів відпустки
* COL\_Excl\_rent: Вартість життя, виключаючи ренту
* COL\_incl\_rent: Вартість життя, включаючи ренту
* Pur\_Power\_Gross: Купівельна спроможність до утримання податків
* Pur\_Power\_Net: Купівельна спроможність після утримання податків
* Pur\_Power\_Annual: Середня купівельна спроможність за рік
* BigMac: Кількість хвилин роботи, необхідних для покупки БігМака
* Bread.kg\_in\_min.: Кількість хвилин роботи, необхідних для покупки кіло хліба
* Rice.kg\_in\_min.: Кількість хвилин роботи, необхідних для покупки кіло риса
* Goods\_and\_Services: Вартість послуг
* Cood\_and\_Services\_Index: Індекс вартості послуг.
* Food\_Index: Індекс вартості їжі.

**Завдання.**

* Побудувати матричну діаграму розсіювання даних і на його основі зробити аналіз мультиколінеарності в моделі.
* Побудувати кореляційну матрицю змінних і кольорову кореляційну матрицю. Зробити висновки.
* Побудувати модель множинної регресії залежності змінної COL\_incl\_rent від решти змінних. Оцінити якість моделі по всіх показниках. За критерієм Фішера з рівнем значущості перевірити гіпотезу ,де – коефіцієнти при відповідних регресорах. Перевірити правильність вашої процедури за допомогою функції linearHypothesis(), (попередньо завантажити пакет car).
* Побудувати лінійну регресійну модель залежності вартості життя COL\_incl\_rent від інших змінних, залучивши до моделі тільки найбільш значущі змінні, проаналізувати залежність саме від них. Для пошуку оптимальної множини регресорів застосувати «регресію вперед», «регресію назад» і комбінацію цих двох підходів. Проаналізувати залишки моделі: predicted versus residuals, нормальність залишків. Оцінити якість моделі по всіх показниках.
* Побудувати лінійну регресійну модель залежності BigMac - вартості БігМак в хвилинах роботи - від інших змінних, залучивши до моделі тільки найбільш значущі змінні, проаналізувати залежність саме від них. Для пошуку оптимальної множини регресорів застосувати «регресію вперед», «регресію назад» і комбінацію цих двох підходів. Проаналізувати залишки моделі: predicted versus residuals, нормальність залишків. Порівняти з моделлю, отриманою в п.4.
* Виписати найбільш оптимальну модель.
* Зробити висновки: які змінні впливають на вартість життя, які – ні? Які змінні слід в будь-якому випадку вилучити з моделі? Чи можна за змінною BigMac оцінювати вартість життя в місті?

**1 Матрична діаграма розсіювання даних та мультиколінеарність**

**library(ggplot2)**

**library(corrplot)**

**library(car)**

**library(olsrr)**

**data<-read.table(file.choose(),header=TRUE)**

**attach(data)**

**pairs(data[2:21],col="darkgreen") - міста з вибірки виключаємо**



Потенційні мультиколінеарні змінні:

1. Wage\_Gross і Wage\_Net:

На графіку ми дуже щільний лінійний зв'язок між цими змінними, що логічно, оскільки чиста зарплата виводиться з валової зарплати після вирахування податків.

2. COL\_Excl\_rent і COL\_incl\_rent:

Ці змінні також мають високий лінійний зв'язок, оскільки вартість життя включає ренту, і тому вони корелюють між собою.

3. Pur\_Power\_Gross і Pur\_Power\_Net:

Ці змінні мають високу кореляцію, оскільки купівельна спроможність до та після податків пов'язана між собою.

4. Food\_Costs і Food\_Index:

Індекс вартості їжі (Food\_Index) базується на вартості харчування (Food\_Costs), що може призводити до високої кореляції між цими змінними.

5. Goods\_and\_Services і Goods\_and\_Services\_Index:

Індекс вартості послуг (Goods\_and\_Services\_Index) може мати високий кореляційний зв'язок з вартістю послуг (Goods\_and\_Services), оскільки індекс базується на вартості.

**2 Кореляційна матриця змінних**

**M<-cor(data[2:21])**

**M**

**Food\_Costs... Womens\_Clothing... Mens\_Clothing...**

**Food\_Costs... 1.0000000 0.6135993 0.6419144**

**Womens\_Clothing... 0.6135993 1.0000000 0.8867571**

**Mens\_Clothing... 0.6419144 0.8867571 1.0000000**

**iPhone\_4S.hr. -0.5891467 -0.5829400 -0.5940529**

**Clothing\_Index 0.6480347 0.9625291 0.9788230**

**Hours\_Worked -0.1385789 -0.3510572 -0.3998805**

**Wage\_Gross 0.7141998 0.6616407 0.6759851**

**Wage\_Net 0.7330766 0.6749564 0.6846065**

**Vacation\_Days -0.0395235 0.2369207 0.1821540**

**COL\_Excl\_rent 0.8651796 0.7218799 0.7530854**

**COL\_incl\_rent 0.8732000 0.6986709 0.7283419**

**Pur\_Power\_Gross 0.5941436 0.5813611 0.5936203**

**Pur\_Power\_Net 0.6010581 0.5903056 0.5952093**

**Pur\_Power\_Annual 0.6407340 0.5676924 0.5621266**

**BigMac -0.4997753 -0.5650307 -0.5563285**

**Bread.kg\_in\_min. -0.1480392 -0.4436233 -0.3500638**

**Rice.kg\_in\_min. -0.5485689 -0.4071119 -0.4253713**

**Goods\_and\_Services... 0.8652695 0.7219425 0.7532138**

**Good\_and\_Services\_Index 0.8651796 0.7218799 0.7530854**

**Food\_Index 0.9999969 0.6132239 0.6418656**

**iPhone\_4S.hr. Clothing\_Index Hours\_Worked**

**Food\_Costs... -0.5891467 0.6480347 -0.1385789**

**Womens\_Clothing... -0.5829400 0.9625291 -0.3510572**

**Mens\_Clothing... -0.5940529 0.9788230 -0.3998805**

**iPhone\_4S.hr. 1.0000000 -0.6057862 0.5217654**

**Clothing\_Index -0.6057862 1.0000000 -0.3882536**

**Hours\_Worked 0.5217654 -0.3882536 1.0000000**

**Wage\_Gross -0.7910181 0.6898407 -0.4899402**

**Wage\_Net -0.8075910 0.7006644 -0.4408331**

**Vacation\_Days -0.2637534 0.2096461 -0.6195768**

**COL\_Excl\_rent -0.7385536 0.7610307 -0.4242037**

**COL\_incl\_rent -0.7343347 0.7362771 -0.3609698**

**Pur\_Power\_Gross -0.8472623 0.6059493 -0.5506458**

**Pur\_Power\_Net -0.8661793 0.6106386 -0.4806232**

**Pur\_Power\_Annual -0.8519786 0.5815756 -0.3280554**

**BigMac 0.9118156 -0.5758002 0.4410907**

**Bread.kg\_in\_min. 0.6621226 -0.4011601 0.3475669**

**Rice.kg\_in\_min. 0.7518201 -0.4293512 0.3488904**

**Goods\_and\_Services... -0.7388191 0.7611307 -0.4244317**

**Good\_and\_Services\_Index -0.7385536 0.7610307 -0.4242037**

**Food\_Index -0.5889204 0.6478387 -0.1384420**

**Wage\_Gross Wage\_Net Vacation\_Days COL\_Excl\_rent**

**Food\_Costs... 0.7141998 0.7330766 -0.03952350 0.8651796**

**Womens\_Clothing... 0.6616407 0.6749564 0.23692073 0.7218799**

**Mens\_Clothing... 0.6759851 0.6846065 0.18215405 0.7530854**

**iPhone\_4S.hr. -0.7910181 -0.8075910 -0.26375339 -0.7385536**

**Clothing\_Index 0.6898407 0.7006644 0.20964612 0.7610307**

**Hours\_Worked -0.4899402 -0.4408331 -0.61957676 -0.4242037**

**Wage\_Gross 1.0000000 0.9816082 0.24252569 0.8989143**

**Wage\_Net 0.9816082 1.0000000 0.21482051 0.8979833**

**Vacation\_Days 0.2425257 0.2148205 1.00000000 0.2252524**

**COL\_Excl\_rent 0.8989143 0.8979833 0.22525245 1.0000000**

**COL\_incl\_rent 0.8794983 0.8896150 0.19417546 0.9917008**

**Pur\_Power\_Gross 0.9586630 0.9398954 0.24906677 0.7838372**

**Pur\_Power\_Net 0.9241943 0.9529363 0.21373504 0.7686785**

**Pur\_Power\_Annual 0.8981658 0.9380779 0.10176696 0.7560826**

**BigMac -0.7301054 -0.7567725 -0.20160911 -0.6369116**

**Bread.kg\_in\_min. -0.5203368 -0.5311508 -0.31703611 -0.3676413**

**Rice.kg\_in\_min. -0.6646503 -0.6950359 -0.10913240 -0.6435915**

**Goods\_and\_Services... 0.8989936 0.8980637 0.22527069 0.9999987**

**Good\_and\_Services\_Index 0.8989143 0.8979833 0.22525245 1.0000000**

**Food\_Index 0.7139138 0.7327020 -0.03951097 0.8651484**

**COL\_incl\_rent Pur\_Power\_Gross Pur\_Power\_Net**

**Food\_Costs... 0.8732000 0.5941436 0.6010581**

**Womens\_Clothing... 0.6986709 0.5813611 0.5903056**

**Mens\_Clothing... 0.7283419 0.5936203 0.5952093**

**iPhone\_4S.hr. -0.7343347 -0.8472623 -0.8661793**

**Clothing\_Index 0.7362771 0.6059493 0.6106386**

**Hours\_Worked -0.3609698 -0.5506458 -0.4806232**

**Wage\_Gross 0.8794983 0.9586630 0.9241943**

**Wage\_Net 0.8896150 0.9398954 0.9529363**

**Vacation\_Days 0.1941755 0.2490668 0.2137350**

**COL\_Excl\_rent 0.9917008 0.7838372 0.7686785**

**COL\_incl\_rent 1.0000000 0.7639613 0.7627355**

**Pur\_Power\_Gross 0.7639613 1.0000000 0.9679291**

**Pur\_Power\_Net 0.7627355 0.9679291 1.0000000**

**Pur\_Power\_Annual 0.7650085 0.9326698 0.9791653**

**BigMac -0.6356412 -0.7913251 -0.8272957**

**Bread.kg\_in\_min. -0.3581337 -0.5832114 -0.6033588**

**Rice.kg\_in\_min. -0.6463409 -0.7011406 -0.7433221**

**Goods\_and\_Services... 0.9917016 0.7839986 0.7688592**

**Good\_and\_Services\_Index 0.9917008 0.7838372 0.7686785**

**Food\_Index 0.8731552 0.5937047 0.6005307**

**Pur\_Power\_Annual BigMac Bread.kg\_in\_min.**

**Food\_Costs... 0.6407340 -0.4997753 -0.1480392**

**Womens\_Clothing... 0.5676924 -0.5650307 -0.4436233**

**Mens\_Clothing... 0.5621266 -0.5563285 -0.3500638**

**iPhone\_4S.hr. -0.8519786 0.9118156 0.6621226**

**Clothing\_Index 0.5815756 -0.5758002 -0.4011601**

**Hours\_Worked -0.3280554 0.4410907 0.3475669**

**Wage\_Gross 0.8981658 -0.7301054 -0.5203368**

**Wage\_Net 0.9380779 -0.7567725 -0.5311508**

**Vacation\_Days 0.1017670 -0.2016091 -0.3170361**

**COL\_Excl\_rent 0.7560826 -0.6369116 -0.3676413**

**COL\_incl\_rent 0.7650085 -0.6356412 -0.3581337**

**Pur\_Power\_Gross 0.9326698 -0.7913251 -0.5832114**

**Pur\_Power\_Net 0.9791653 -0.8272957 -0.6033588**

**Pur\_Power\_Annual 1.0000000 -0.8231854 -0.5883453**

**BigMac -0.8231854 1.0000000 0.6562998**

**Bread.kg\_in\_min. -0.5883453 0.6562998 1.0000000**

**Rice.kg\_in\_min. -0.7464141 0.6822046 0.3664806**

**Goods\_and\_Services... 0.7562486 -0.6372760 -0.3677717**

**Good\_and\_Services\_Index 0.7560826 -0.6369116 -0.3676413**

**Food\_Index 0.6402446 -0.4993592 -0.1477465**

**Rice.kg\_in\_min. Goods\_and\_Services...**

**Food\_Costs... -0.5485689 0.8652695**

**Womens\_Clothing... -0.4071119 0.7219425**

**Mens\_Clothing... -0.4253713 0.7532138**

**iPhone\_4S.hr. 0.7518201 -0.7388191**

**Clothing\_Index -0.4293512 0.7611307**

**Hours\_Worked 0.3488904 -0.4244317**

**Wage\_Gross -0.6646503 0.8989936**

**Wage\_Net -0.6950359 0.8980637**

**Vacation\_Days -0.1091324 0.2252707**

**COL\_Excl\_rent -0.6435915 0.9999987**

**COL\_incl\_rent -0.6463409 0.9917016**

**Pur\_Power\_Gross -0.7011406 0.7839986**

**Pur\_Power\_Net -0.7433221 0.7688592**

**Pur\_Power\_Annual -0.7464141 0.7562486**

**BigMac 0.6822046 -0.6372760**

**Bread.kg\_in\_min. 0.3664806 -0.3677717**

**Rice.kg\_in\_min. 1.0000000 -0.6437158**

**Goods\_and\_Services... -0.6437158 1.0000000**

**Good\_and\_Services\_Index -0.6435915 0.9999987**

**Food\_Index -0.5485795 0.8652386**

**Good\_and\_Services\_Index Food\_Index**

**Food\_Costs... 0.8651796 0.99999688**

**Womens\_Clothing... 0.7218799 0.61322392**

**Mens\_Clothing... 0.7530854 0.64186562**

**iPhone\_4S.hr. -0.7385536 -0.58892038**

**Clothing\_Index 0.7610307 0.64783872**

**Hours\_Worked -0.4242037 -0.13844196**

**Wage\_Gross 0.8989143 0.71391381**

**Wage\_Net 0.8979833 0.73270197**

**Vacation\_Days 0.2252524 -0.03951097**

**COL\_Excl\_rent 1.0000000 0.86514836**

**COL\_incl\_rent 0.9917008 0.87315523**

**Pur\_Power\_Gross 0.7838372 0.59370466**

**Pur\_Power\_Net 0.7686785 0.60053075**

**Pur\_Power\_Annual 0.7560826 0.64024460**

**BigMac -0.6369116 -0.49935916**

**Bread.kg\_in\_min. -0.3676413 -0.14774652**

**Rice.kg\_in\_min. -0.6435915 -0.54857953**

**Goods\_and\_Services... 0.9999987 0.86523865**

**Good\_and\_Services\_Index 1.0000000 0.86514836**

**Food\_Index 0.8651484 1.00000000**

**2.1) Кольорова кореляційна матриця даних**

**corrplot(M,method="color")**



**Vacation days майже не корелює**

**Iphone 4s має зворотню кореляцію як і індекс BigMack**

**3 Модель множинної регресії COL\_incl\_rent від решти змінних.**

**COLIR<-lm(COL\_incl\_rent ~ Food\_Costs... + Womens\_Clothing... +Mens\_Clothing... +iPhone\_4S.hr. +Clothing\_Index +Hours\_Worked +Wage\_Gross+Wage\_Net+Vacation\_Days +COL\_Excl\_rent +Pur\_Power\_Gross +Pur\_Power\_Net +Pur\_Power\_Annual +BigMac +Bread.kg\_in\_min. +Rice.kg\_in\_min.+ Goods\_and\_Services... +Food\_Index,data=data)**

**summary(COLIR)**

**Call:**

**lm(formula = COL\_incl\_rent ~ Food\_Costs... + Womens\_Clothing... +**

**Mens\_Clothing... + iPhone\_4S.hr. + Clothing\_Index + Hours\_Worked +**

**Wage\_Gross + Wage\_Net + Vacation\_Days + COL\_Excl\_rent + Pur\_Power\_Gross +**

**Pur\_Power\_Net + Pur\_Power\_Annual + BigMac + Bread.kg\_in\_min. +**

**Rice.kg\_in\_min. + Goods\_and\_Services... + Food\_Index, data = data)**

**Residuals:**

**Min 1Q Median 3Q Max**

**-3.4041 -1.0208 -0.2754 0.7117 4.3833**

**Coefficients:**

**Estimate Std. Error t value Pr(>|t|)**

**(Intercept) -25.242164 7.722937 -3.268 0.0019 \*\***

**Food\_Costs... 0.165000 0.797877 0.207 0.8370**

**Womens\_Clothing... 0.004437 0.051275 0.087 0.9314**

**Mens\_Clothing... 0.006244 0.051808 0.121 0.9045**

**iPhone\_4S.hr. 0.001839 0.008957 0.205 0.8381**

**Clothing\_Index -0.101201 0.799312 -0.127 0.8997**

**Hours\_Worked 0.006329 0.002839 2.229 0.0301 \***

**Wage\_Gross -0.438563 0.205407 -2.135 0.0374 \***

**Wage\_Net 0.351224 0.217198 1.617 0.1118**

**Vacation\_Days 0.066977 0.048144 1.391 0.1700**

**COL\_Excl\_rent -8.519805 8.414690 -1.012 0.3159**

**Pur\_Power\_Gross 0.274754 0.168386 1.632 0.1087**

**Pur\_Power\_Net -0.331147 0.181714 -1.822 0.0740 .**

**Pur\_Power\_Annual 0.153910 0.086528 1.779 0.0810 .**

**BigMac -0.007017 0.032125 -0.218 0.8279**

**Bread.kg\_in\_min. 0.008757 0.031944 0.274 0.7850**

**Rice.kg\_in\_min. 0.021106 0.045591 0.463 0.6453**

**Goods\_and\_Services... 0.244831 0.213974 1.144 0.2577**

**Food\_Index -0.949073 4.402816 -0.216 0.8302**

**---**

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

**Residual standard error: 1.801 on 53 degrees of freedom**

**Multiple R-squared: 0.9923, Adjusted R-squared: 0.9897**

**F-statistic: 381.3 on 18 and 53 DF, p-value: < 2.2e-16**

**alias - Good\_and\_Services\_Index спричиняє мультиколінеарність**

**> alias(COLIR)**

**Model :**

**COL\_incl\_rent ~ Food\_Costs... + Womens\_Clothing... + Mens\_Clothing... +**

**iPhone\_4S.hr. + Clothing\_Index + Hours\_Worked + Wage\_Gross +**

**Wage\_Net + Vacation\_Days + COL\_Excl\_rent + Pur\_Power\_Gross +**

**Pur\_Power\_Net + Pur\_Power\_Annual + BigMac + Bread.kg\_in\_min. +**

**Rice.kg\_in\_min. + Goods\_and\_Services... + Good\_and\_Services\_Index +**

**Food\_Index**

**Complete :**

**(Intercept) Food\_Costs... Womens\_Clothing... Mens\_Clothing... iPhone\_4S.hr. Clothing\_Index Hours\_Worked Wage\_Gross**

**Good\_and\_Services\_Index 0 0 0 0 0 0 0 0**

**Wage\_Net Vacation\_Days COL\_Excl\_rent Pur\_Power\_Gross Pur\_Power\_Net Pur\_Power\_Annual BigMac Bread.kg\_in\_min.**

**Good\_and\_Services\_Index 0 0 1 0 0 0 0 0**

**Rice.kg\_in\_min. Goods\_and\_Services... Food\_Index**

**Good\_and\_Services\_Index 0 0 0**

**3.1 Оцінка за критерієм Фішера**

Перевірка гіпотези H\_0={β\_ = 1.2 \*β\_}

hypothesis <- " - 1.2 \* = 0"

linearHypothesis(COLIR, hypothesis)

**Linear hypothesis test**

**Hypothesis:**

**Pur\_Power\_Gross - 1.2 Pur\_Power\_Net = 0**

**Model 1: restricted model**

**Model 2: COL\_incl\_rent ~ Food\_Costs... + Womens\_Clothing... + Mens\_Clothing... +**

**iPhone\_4S.hr. + Clothing\_Index + Hours\_Worked + Wage\_Gross +**

**Wage\_Net + Vacation\_Days + COL\_Excl\_rent + Pur\_Power\_Gross +**

**Pur\_Power\_Net + Pur\_Power\_Annual + BigMac + Bread.kg\_in\_min. +**

**Rice.kg\_in\_min. + Goods\_and\_Services... + Food\_Index**

**Res.Df RSS Df Sum of Sq F Pr(>F)**

**1 54 182.28**

**2 53 171.83 1 10.457 3.2253 0.07821 .**

**---**

**За критерієм Фішера з рівнем значущості α=0.05, оскільки отримане p-значення (0.07821) більше ніж 0.05.**

**4 Лінійна регресійна модель COL\_incl\_rent від найбільш значущі змінних**

**COLINSA<-lm(COL\_incl\_rent ~ Hours\_Worked +Wage\_Gross +Wage\_Net+Vacation\_Days+COL\_Excl\_rent +Pur\_Power\_Gross+Bread.kg\_in\_min. +Rice.kg\_in\_min. ,data=data)**

**4.1 Регрессія назад**

backward\_model <- ols\_step\_backward\_p(COLINSA)

summary(backward\_model$model)

Call:

lm(formula = paste(response, "~", paste(c(include, cterms), collapse = " + ")),

data = l)

Residuals:

Min 1Q Median 3Q Max

-3.2339 -1.2310 -0.0279 0.8967 5.9987

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -28.012746 5.447000 -5.143 2.69e-06 \*\*\*

Hours\_Worked 0.008531 0.001860 4.586 2.11e-05 \*\*\*

Wage\_Gross -0.203426 0.048831 -4.166 9.34e-05 \*\*\*

Wage\_Net 0.100768 0.039589 2.545 0.0133 \*

Vacation\_Days 0.057757 0.043647 1.323 0.1904

COL\_Excl\_rent 1.050056 0.036553 28.727 < 2e-16 \*\*\*

Pur\_Power\_Gross 0.089324 0.034610 2.581 0.0121 \*

---

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

Residual standard error: 1.805 on 65 degrees of freedom

Multiple R-squared: 0.9906, Adjusted R-squared: 0.9897

F-statistic: 1137 on 6 and 65 DF, p-value: < 2.2e-16

**4.2 Регрессія вперед**

forward\_model <- ols\_step\_forward\_p(COLINSA)

summary(forward\_model$model)

Call:

lm(formula = paste(response, "~", paste(preds, collapse = " + ")),

data = l)

Residuals:

Min 1Q Median 3Q Max

-3.2009 -1.3407 -0.0190 0.9826 6.0143

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -28.941095 5.977149 -4.842 8.49e-06 \*\*\*

COL\_Excl\_rent 1.056540 0.040375 26.168 < 2e-16 \*\*\*

Hours\_Worked 0.008582 0.001877 4.572 2.27e-05 \*\*\*

Rice.kg\_in\_min. 0.015756 0.040394 0.390 0.697784

Wage\_Gross -0.213465 0.055483 -3.847 0.000278 \*\*\*

Wage\_Net 0.104118 0.040765 2.554 0.013037 \*

Pur\_Power\_Gross 0.096755 0.039706 2.437 0.017611 \*

Vacation\_Days 0.056978 0.043980 1.296 0.199780

---

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

Residual standard error: 1.817 on 64 degrees of freedom

Multiple R-squared: 0.9906, Adjusted R-squared: 0.9896

F-statistic: 961.6 on 7 and 64 DF, p-value: < 2.2e-16

**4.3 Регресія комбінована**

stepwise\_model <- ols\_step\_both\_p(COLINSA)

summary(stepwise\_model$model)

Call:

lm(formula = paste(response, "~", paste(preds, collapse = " + ")),

data = l)

Residuals:

Min 1Q Median 3Q Max

-3.6662 -1.0624 -0.2568 0.9344 6.0060

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -18.624869 3.199612 -5.821 1.68e-07 \*\*\*

COL\_Excl\_rent 0.985274 0.014137 69.694 < 2e-16 \*\*\*

Hours\_Worked 0.006879 0.001386 4.963 4.81e-06 \*\*\*

---

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

Residual standard error: 1.99 on 69 degrees of freedom

Multiple R-squared: 0.9878, Adjusted R-squared: 0.9875

F-statistic: 2798 on 2 and 69 DF, p-value: < 2.2e-16

**4.4 Predicted vs residuals, нормальність залишків**

qqnorm(COLINSA$residuals)

qqline(COLINSA$residuals, col=5)

plot(COLINSA$fitted.values,COLINSA$residuals, col=3)



**5 Лінійна регресійна модель BigMac від найбільш значущих**

**змінних**

**MAC<-lm(BigMac ~ Hours\_Worked +Wage\_Gross +Wage\_Net+Vacation\_Days+COL\_Excl\_rent +Pur\_Power\_Gross+Bread.kg\_in\_min. +Rice.kg\_in\_min. ,data=data)**

**5.1 Регрессія назад**

backward\_model <- ols\_step\_backward\_p(MAC)

summary(backward\_model$model)

Call:

lm(formula = paste(response, "~", paste(c(include, cterms), collapse = " + ")),

data = l)

Residuals:

Min 1Q Median 3Q Max

-18.492 -3.873 -0.665 4.325 31.630

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 65.3368 15.6844 4.166 9.34e-05 \*\*\*

Wage\_Gross 0.9074 0.2901 3.128 0.002637 \*\*

Wage\_Net -0.3903 0.2072 -1.884 0.064005 .

COL\_Excl\_rent -0.4394 0.1999 -2.199 0.031476 \*

Pur\_Power\_Gross -0.6814 0.1905 -3.578 0.000661 \*\*\*

Bread.kg\_in\_min. 0.4497 0.1235 3.642 0.000537 \*\*\*

Rice.kg\_in\_min. 0.2608 0.2100 1.242 0.218708

---

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

Residual standard error: 9.472 on 65 degrees of freedom

Multiple R-squared: 0.7596, Adjusted R-squared: 0.7374

F-statistic: 34.22 on 6 and 65 DF, p-value: < 2.2e-16

**5.2 Регрессія вперед**

forward\_model <- ols\_step\_forward\_p(MAC)

summary(forward\_model$model)

Call:

lm(formula = paste(response, "~", paste(preds, collapse = " + ")),

data = l)

Residuals:

Min 1Q Median 3Q Max

-21.4631 -5.2530 -0.3573 5.1703 31.5566

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 26.58238 7.49065 3.549 0.000707 \*\*\*

Pur\_Power\_Gross -0.24186 0.06001 -4.031 0.000143 \*\*\*

Bread.kg\_in\_min. 0.49843 0.12559 3.969 0.000177 \*\*\*

Rice.kg\_in\_min. 0.58747 0.19079 3.079 0.002993 \*\*

---

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

Residual standard error: 9.95 on 68 degrees of freedom

Multiple R-squared: 0.7224, Adjusted R-squared: 0.7102

F-statistic: 58.99 on 3 and 68 DF, p-value: < 2.2e-16

**5.3 Регресія комбінована**

stepwise\_model <- ols\_step\_both\_p(MAC)

summary(stepwise\_model$model)

Call:

lm(formula = paste(response, "~", paste(preds, collapse = " + ")),

data = l)

Residuals:

Min 1Q Median 3Q Max

-21.4631 -5.2530 -0.3573 5.1703 31.5566

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 26.58238 7.49065 3.549 0.000707 \*\*\*

Pur\_Power\_Gross -0.24186 0.06001 -4.031 0.000143 \*\*\*

Bread.kg\_in\_min. 0.49843 0.12559 3.969 0.000177 \*\*\*

Rice.kg\_in\_min. 0.58747 0.19079 3.079 0.002993 \*\*

---

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

Residual standard error: 9.95 on 68 degrees of freedom

Multiple R-squared: 0.7224, Adjusted R-squared: 0.7102

F-statistic: 58.99 on 3 and 68 DF, p-value: < 2.2e-16

**5.4 Predicted vs residuals, нормальність залишків**

qqnorm(MAC$residuals)

qqline(MAC$residuals, col=5)

plot(MAC$fitted.values,MAC$residuals, col=3)





**6) Найбільш оптимальна модель**

Найбільш оптимальна модель є Backward\_model

**7)Висновки**

На основі аналізу регресійних моделей :

Змінні, які впливають на вартість життя:

1. Hours\_Worked Статистично значущий коефіцієнт вказує на те, що середня кількість робочих годин впливає на вартість життя. Це може бути пов'язано з рівнем доходів та економічною активністю.

2. Wage\_Gross: Статистично значущий коефіцієнт показує, що середня зарплатня до утримання податків також має великий вплив на вартість життя. Це пов'язано з доступністю коштів для витрат.

3. Wage\_Net: Хоча ця змінна не досягла статистичної значущості на рівні 0.05, вона може мати певний вплив на вартість життя.

Змінні, які, ймовірно, не впливають на вартість життя.

iPhone\_4S(hr): Ця змінна також не є статистично значущою, що означає, що вартість iPhone 4S не має великого впливу на вартість життя.

Вилучити з моделі слід City, Vacation days та Good\_and\_Services\_Index

BigMac може бути корисним для порівняння вартості життя між різними містами або для оцінки змін вартості життя в одному місті з часом, але він не враховує всі аспекти вартості життя, такі як вартість житла, освіти, охорони здоров'я і т. д.

Оскільки BigMac є показником вартості життя, який базується на цінах на продукти харчування, він може відображати лише частину загальної вартості життя і не може давати повний обсяг інформації про життєві витрати.

Вартість BigMac може змінюватися в залежності від локальних факторів, таких як локальна економіка, рівень заробітної плати, валютні курси і т. д. Тому важливо враховувати ці фактори при використанні BigMac для оцінки вартості життя.