

## PROJECT UJIAN AKHIR SEMESTER

### MATA KULIAH MODEL LINEAR

Judul : Uji Asumsi Klasik Regresi Linear Berganda Pengaruh “rank”, “yrs.service”, dan “sex” terhadap “salary”.

Nama : Rosa Amalia Nursinta

NIM : 11190940000041

Kelas : Matematika 5B

Dosen Pengajar : Ary Santoso, S.Stat, M.Si, M.I.Pol

#### Data yang akan diolah :

Akan dilakukan analisis data sample “Salaries\_sample” sebanyak 200 amatan dari data populasi “Salaries” yang diambil menggunakan fungsi “sample\_n” dengan set.seed (41) dengan pengujian asumsi klasik secara visual (deskriptif) maupun inferensia untuk memperoleh model terbaik dalam melihat hubungan sebab akibat antara peubah “rank”, “yrs.service”, dan “sex” terhadap “salary” menggunakan software R dengan tindakan mengatasinya jika terjadi pelanggaran asumsi.

#### Langkah analisis :

1. Mengambil Data Sample : Diambil data sample sebanyak 200 amatan dari data populasi menggunakan fungsi “sample\_n” dengan set.seed (41).
2. Membentuk Variabel Dummy : Mengubah nilai data kategorik menjadi bilangan binary 1,0.
3. Menentukan Model Regresi : Membentuk model regresi dari 200 data sample yang telah diambil.
4. Uji Signifikansi Persamaan Regresi (Uji simultan F) : Uji ini digunakan untuk melihat apakah variabel independen secara bersama-sama mempengaruhi variabel dependen.
5. Uji Signifikansi Koefisien Regresi (Uji parsial t) : Uji ini digunakan untuk melihat apakah tiap-tiap variabel independen memberikan pengaruh yang signifikan terhadap variabel dependen atau tidak. Model yang baik adalah model yang tiap variabel independen memberi pengaruh yang signifikan terhadap variabel dependen.
6. Uji Normalitas : Uji ini digunakan untuk melihat apakah data berdistribusi normal atau tidak. Model yang baik adalah model yang residual nya berdistribusi normal.
7. Uji Heterokedastisitas : Uji ini digunakan untuk melihat apakah variansi dari residual homogen atau tidak. Model yang baik adalah model dengan data yang variansi dari residualnya bersifat homogen.
8. Uji Multikolinearitas : Uji ini digunakan untuk melihat apakah antar variabel bebas memiliki hubungan atau berkorelasi tinggi. Model yang baik adalah model yang antar variabel bebas nya tidak memiliki korelasi yang tinggi.
9. Uji Autokorelasi : Uji ini digunakan untuk melihat apakah data observasi memiliki korelasi dengan data observasi sebelumnya. Model yang baik adalah model dengan data observasi tidak memiliki korelasi dengan data sebelumnya.
10. Melakukan perbaikan model jika terjadi pelanggaran terhadap asumsi yang diuji.

11. Pengujian asumsi kembali model regresi yang telah diperbaiki.
12. Kesimpulan

## Pembahasan dan Interpretasi

### 1. Pengambilan Data Sample

Diambil data sample "Salaries\_sample" sebanyak 200 amatan dari data populasi "Salaries" menggunakan fungsi "sample\_n" dengan set.seed (41), sehingga diperoleh data sebagai berikut :

```
> # Memasukkan set data populasi
> Salaries <- read.csv("D:/Salaries.csv")
> View(Salaries)
>
> # Mengambil 200 data sampel
> data("Salaries", package="car")
Warning message:
In data("Salaries", package = "car") : data set 'Salaries' not found
> set.seed(41)
> Salaries_sample=sample_n(Salaries,200)
> View(Salaries_sample)
>
> # Melihat Keragaman data Rank dan Gender
> table(Salaries_sample$rank)

AssocProf  AsstProf      Prof
        35         33        132
> table(Salaries_sample$sex)

Female   Male
      22    178
> table(Salaries_sample$rank,Salaries_sample$sex)

          Female Male
AssocProf      5   30
AsstProf       6   27
Prof          11  121
>
```

#### Interpretasi :

Sebanyak 200 data yang telah diambil, kemudian data sample ini dinamakan sebagai "Salaries\_Sample".

Berdasarkan 200 data ini, dapat dilihat bahwa keragaman data "rank" dan "sex" yaitu sebanyak 35 staff bekerja sebagai Associate Prof, 33 staff sebagai Assistant Prof, dan sebanyak 132 staff adalah Professor. Kemudian di antara 200 staff tersebut sebanyak 22 staff adalah perempuan dan sebanyak 178 staff adalah laki-laki.

Kemudian di antara 35 Associate Prof tersebut, sebanyak 30 orang adalah Associate Prof laki-laki dan sebanyak 5 orang adalah Associate Prof perempuan. Di antara 33 Assistant Professor, sebanyak 27 orang adalah Assistant Prof laki-laki dan sebanyak 6 orang adalah Assistant Prof perempuan. Dan di antara 132 professor, sebanyak 121 orang adalah professor laki-laki dan sebanyak 11 orang adalah professor perempuan.

## 2. Membentuk Variabel Dummy

Peubah rank dikategorikan dalam 3 kategori, yaitu associate professor (AssocProf), assistant professor (AsstProf), dan professor (Prof). Karena rank terbagi dalam 3 kategori ( $k=3$ ), maka peubah dummy nya  $k-1$  yaitu 2 peubah dummy, yaitu dAsstP dan dProf. Kemudian untuk peubah sex dikategorikan dalam 2 kategori, maka peubah dummy yang dibentuk adalah 1 peubah dummy yaitu dMale.

```
> # Mengubah nilai kategorik rank menjadi nilai bilangan binary 1,0
>
> # Memisalkan rank AssocProf=0, AsstProf=1, Prof=1
> dAsstP<-ifelse(Salaries_sample$rank=='AsstProf',1,0)
> dProf<-ifelse(Salaries_sample$rank=='Prof',1,0)
>
> # Memisalkan sex Male=1, Female=0
> dMale<-ifelse(Salaries_sample$sex=='Male',1,0)
>
> # Menyatukan data rank, sex, yrs.service, dan salary
> rank2<-cbind(dAsstP,dProf)
> rank_gender<-cbind(rank2,dMale)
> Salaries2<-cbind(Salaries_sample$salary,Salaries_sample$yrs.service,rank_gender)
> Salaries_41<-data.frame(Salaries2)
> view(Salaries_41)
```

### Interpretasi :

Variabel dummy yang dibuat untuk peubah rank adalah memisalkan ketegori AssocProf=0, AssistProf=1, dan Prof=1. Kemudian variabel dummy untuk peubah sex adalah memisalkan kategori Male=1 dan Female=0.

Kemudian peubah dummy dAsstProf, dProf, dan dMale akan disatukan ke suatu data set baru beserta peubah yrs.service dan salary yang dinamakan sebagai data “Salaries\_41”.

## 3. Membentuk Model Regresi

Membentuk model regresi dari 200 data sample yang telah diambil menggunakan peubah dummy yang telah dibuat. Untuk mengetahui pengaruh rank, sex, yrs.service terhadap salary maka model regresi yang dibentuk adalah sebagai berikut :

```
> # Memodelkan Pengaruh rank, sex, yrs.service terhadap salary
> modelSalaries_41<-lm(formula=V1~V2+dAsstP+dProf+dMale, data=Salaries_41)
> summary(modelSalaries_41)

Call:
lm(formula = V1 ~ V2 + dAsstP + dProf + dMale, data = Salaries_41)

Residuals:
    Min       1Q   Median       3Q      Max
-60507 -14078  -1087   10803   85123

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  91369.0     6330.2   14.434 < 2e-16 ***
V2           -258.8       155.8   -1.661  0.0982 .
dAsstP       -13343.0     5927.9   -2.251  0.0255 *
dProf        35190.9     4789.0    7.348 5.37e-12 ***
dMale        4946.8      5385.1    0.919  0.3594
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23610 on 195 degrees of freedom
Multiple R-squared:  0.3885, Adjusted R-squared:  0.376
F-statistic: 30.97 on 4 and 195 DF, p-value: < 2.2e-16

>
```

### Model Regresi :

- Model regresi keseluruhan :  $V1 = 91369.0 - 258.8 V2 - 13343.0 dAsstP + 35190.9 dProf + 4946.8 dMale + e$   
Dengan  $V1 = \text{Salary}$ ,  $V2 = \text{yrs.service}$
- Model untuk rank Professor laki-laki ( $Prof=1, AsstProf=0, Male=1, Female=0$ )  
 $\text{Salary} = 91369.0 - 258.8 \text{ yrs.service} + 35190.9 + 4946.8 + e$   
 $\text{Salary} = 131506.7 - 258.8 \text{ yrs.service} + e$
- Model untuk rank Professor perempuan ( $Prof=1, AsstProf=0, Male=0, Female=1$ )  
 $\text{Salary} = 91369.0 - 258.8 \text{ yrs.service} + 35190.9 + e$   
 $\text{Salary} = 126559.9 - 258.8 \text{ yrs.service} + e$
- Model untuk rank Assistant Professor laki-laki ( $Prof=0, AsstProf=1, Male=1, Female=0$ )  
 $\text{Salary} = 91369.0 - 258.8 \text{ yrs.service} - 13343.0 + 4946.8 + e$   
 $\text{Salary} = 82972.8 - 258.8 \text{ yrs.service} + e$
- Model untuk rank Assistant Professor perempuan ( $Prof=0, AsstProf=1, Male=0, Female=1$ )  
 $\text{Salary} = 91369.0 - 258.8 \text{ yrs.service} - 13343.0 + e$   
 $\text{Salary} = 78026 - 258.8 \text{ yrs.service} + e$
- Model untuk rank Associate Professor laki-laki ( $Prof=1, AsstProf=0, Male=1, Female=0$ )  
 $\text{Salary} = 91369.0 - 258.8 \text{ yrs.service} + 4946.8 + e$   
 $\text{Salary} = 140837 - 258.8 \text{ yrs.service} + e$
- Model untuk rank Associate Professor perempuan ( $Prof=1, AsstProf=0, Male=0, Female=1$ )  
 $\text{Salary} = 91369.0 - 258.8 \text{ yrs.service} + e$

### Interpretasi :

- Berdasarkan output yang dihasilkan didapat nilai koefisien  $b_0, b_1, b_2, b_3, b_4$  berturut-turut adalah 91369.0, - 258.8, -13343.0, 35190.9, 4946.8
- Nilai dugaan intercept = 91369.0, artinya rata-rata gaji associate professor perempuan sebesar 91369.0 ketika lama kerja=0.
- Nilai dugaan yrs.service = - 258.8, artinya rata-rata perubahan gaji tiap staff sebesar 258.8 ketika terjadi perubahan 1 tahun lama kerja.
- Nilai dugaan dAsstP = -13343.0, artinya rata-rata perbedaan gaji assistant professor dengan associate professor untuk gender yang sama adalah 13343.0 ketika lama kerja=0.
- Nilai dugaan dProf = 35190.9, artinya rata-rata perbedaan gaji professor dengan associate professor untuk gender yang sama adalah 35190.9 ketika lama kerja=0.
- Nilai dugaan dMale = 4946.8, artinya rata-rata perbedaan gaji staff laki-laki dan perempuan untuk level yang sebanding adalah 4946.8 ketika lama kerja=0.
- Adjusted R-squared : 0.376, artinya sebesar 37.6% model mampu menjelaskan keragaman variabel salary dan sekitar 62.4% keragaman yang tidak mampu dijelaskan oleh variabel penjelas dan terletak pada komponen error.

#### 4. Uji-F Simultan

Uji ini digunakan untuk melihat apakah variabel dProf, dAsstP, dMale, dan yrs.service secara bersama-sama mempengaruhi variabel Salary.

- Hipotesis Uji :  
 $H_0 : \beta_i = 0$  (model tidak layak),  $i = 0,1,2,3,4$   
 $H_1 : \beta_i \neq 0$  (model layak),  $i = 0,1,2,3,4$
- $\alpha: 5\% = 0.05$
- Statistik Uji :

```
Residual standard error: 23610 on 195 degrees of freedom  
Multiple R-squared: 0.3885, Adjusted R-squared: 0.376  
F-statistic: 30.97 on 4 and 195 DF, p-value: < 2.2e-16
```

$$F_{4,195,\alpha=0.05} = 2.42, F_{hitung} = 30.97 \text{ Pvalue} < 2.2e - 16$$

- Daerah Kritis :  
 $H_0$  ditolak jika  $F_{hitung} > F_{tabel}$  atau  $Pvalue < 0.05$
- Keputusan :  
 $F_{4,195,\alpha=0.05} = 2.42, F_{hitung} = 30.97 \text{ Pvalue} < 2.2e - 16$   
 $F_{hitung} > F_{tabel}$  dan  $Pvalue < 0.05$ , maka  $H_0$  ditolak
- Kesimpulan :  
Berdasarkan Uji-F memperlihatkan bahwa cukup bukti untuk variabel dProf, dAsstP, dMale, dan yrs.service secara bersama-sama mempengaruhi variabel Salary.

#### 5. Uji-t Parsial

Uji ini digunakan untuk melihat apakah tiap-tiap variabel dProf, dAsstP, dMale, dan yrs.service berpengaruh signifikan terhadap variabel Salary.

- Hipotesis Uji :  
 $H_0 : \beta_i = 0$  (variabel tidak berpengaruh signifikan),  $i = 0,1,2,3,4$   
 $H_1 : \beta_i \neq 0$  (variabel berpengaruh signifikan),  $i = 0,1,2,3,4$
- $\alpha: 5\% = 0.05$
- Statistik Uji :

```
Coefficients:  
Estimate Std. Error t value Pr(>|t|)  
(Intercept) 91369.0 6330.2 14.434 < 2e-16 ***  
v2 -258.8 155.8 -1.661 0.0982 .  
dAsstP -13343.0 5927.9 -2.251 0.0255 *  
dProf 35190.9 4789.0 7.348 5.37e-12 ***  
dMale 4946.8 5385.1 0.919 0.3594  
---  
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$Sb_1 = 155.8, \quad Sb_2 = 5927.9, \quad Sb_3 = 4789.0, \quad Sb_4 = 5385.1, \\ t_{hit1} = -1.661, \quad t_{hit2} = -2.251, \quad t_{hit3} = 7.348, \quad t_{hit4} = 0.919, \\ t_{195,\alpha=0.05} = 1.97220$$

- Daerah Kritis :

$H_0$  ditolak jika  $-t_{tabel} > t_{hitung} > t_{tabel}$

- Keputusan :

- $t_{195, \alpha=0.05} = 1.97220$

$t_{hit1} = -1.661, t_{hitung} > -t_{tabel}, \text{ maka } H_0 \text{ diterima}$

$t_{hit2} = -2.251, t_{hitung} < -t_{tabel}, \text{ maka } H_0 \text{ ditolak}$

$t_{hit3} = 7.348, t_{hitung} > t_{tabel}, \text{ maka } H_0 \text{ ditolak}$

$t_{hit4} = 0.919, t_{hitung} < t_{tabel}, \text{ maka } H_0 \text{ diterima}$

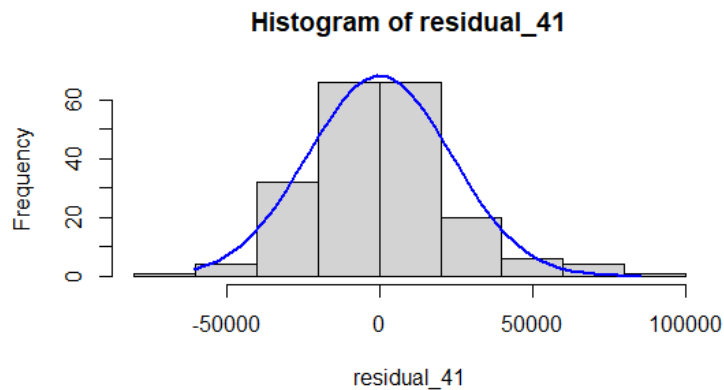
- Kesimpulan :

Berdasarkan Uji-t memperlihatkan bahwa dari serangkaian variabel independent, yang berpengaruh signifikan terhadap Salary adalah dAsstP dan dProf.

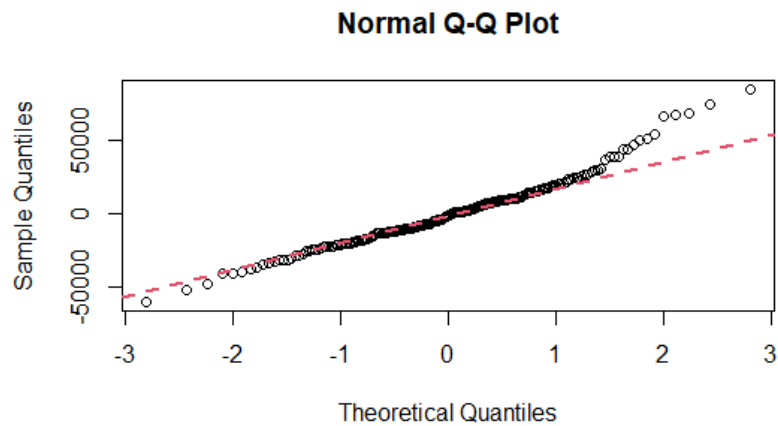
## 6. Uji Normalitas

Uji ini digunakan untuk melihat apakah data berdistribusi normal atau tidak.

- Melihat grafik histogram dan normal Q-QPlot



Secara visual terlihat bahwa garis fit pada grafik histogram ini menyerupai histogram yang berdistribusi normal. Maka secara objektif dapat diasumsikan bahwa residual berdistribusi normal



Secara visual terlihat bahwa penyebaran residual relatif mengikuti garis lurus, ini mengindikasikan bahwa residual berdistribusi normal.

Namun untuk memastikan apakah residual berdistribusi normal atau tidak, diperlukan analisis infrensia menggunakan hipotesis uji

- Hipotesis Uji :  
 $H_0$  : residual berdistribusi normal  
 $H_1$  : residual tidak berdistribusi normal
- $\alpha: 5\% = 0.05$
- Statistik Uji :

```
> # Uji Normalitas
> shapiro.test(residual_41) # terjadi pelanggaran

      Shapiro-Wilk normality test

data:  residual_41
W = 0.96828, p-value = 0.0001743
>
```

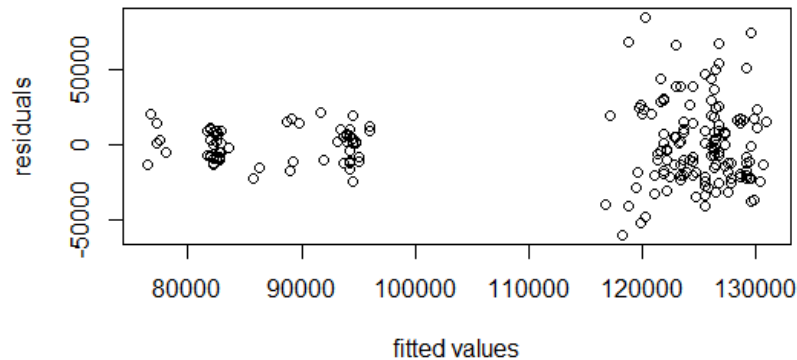
$Pvalue = 0.0001743$

- Daerah Kritis :  
 $H_0$  ditolak jika  $Pvalue < 0.05$
- Keputusan :  
 Berdasarkan hasil pengujian terlihat bahwa  $Pvalue = 0.0001743 < 0.05$  maka  $H_0$  ditolak. Ini berarti bahwa residual tidak berdistribusi normal.

## 7. Uji Heterokedastisitas

Uji ini digunakan untuk melihat apakah variansi dari residual homogen atau tidak.

- Melihat plot Residuals vs Fitted Values



Berdasarkan plot Residual vs Fitted Values terlihat bahwa residual tersebar secara tidak merata maka dapat diasumsikan bahwa variansi dari residual tidak bersifat homogen atau dapat dikatakan terjadi heteroskedastisitas. Selanjutnya kita lakukan analisis inferensia untuk memastikan asumsi heteroskedastisitas menggunakan hipotesis uji

- Hipotesis Uji :  
 $H_0 : \sigma_i^2 = \sigma^2 \ i = 0,1,2,3,4$  (tidak terjadi heteroskedastisitas)  
 $H_1 : \exists \sigma_i^2 \neq \sigma^2 \ i = 0,1,2,3,4$  (terjadi heteroskedastisitas)
- $\alpha: 5\% = 0.05$
- Statistik Uji :

```
> # Uji heterokedastisitas
> bptest(modelSalaries_41, studentize = F) # terjadi pelanggaran

Breusch-Pagan test

data: modelSalaries_41
BP = 48.461, df = 4, p-value = 7.563e-10
>
```

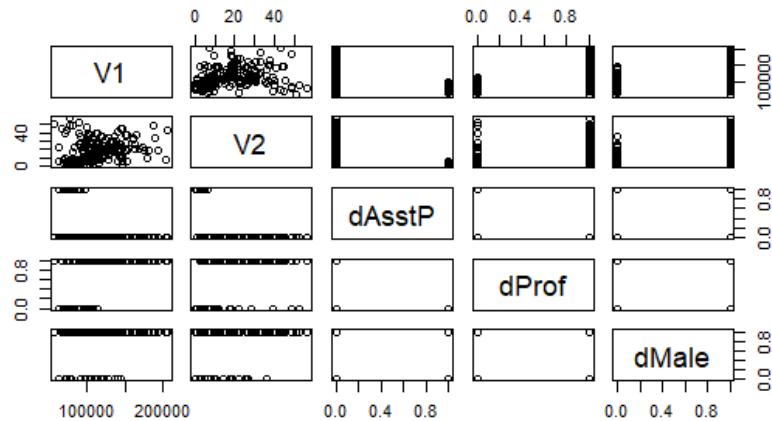
$Pvalue = 7.563e - 10$

- Daerah Kritis :  
 $H_0$  ditolak jika  $Pvalue < 0.05$
- Keputusan :  
Berdasarkan hasil pengujian terlihat bahwa  $Pvalue = 7.563e - 10 < 0.05$  maka  $H_0$  ditolak. Ini berarti bahwa terjadi heteroskedastisitas pada model.



## 8. Uji Multikolinearitas

Uji ini digunakan untuk melihat apakah antar variabel bebas memiliki hubungan atau berkorelasi tinggi.



Secara visual, dapat dilihat bahwa tidak terdapat hubungan linear yang relatif kuat antar variabel bebas, sehingga dapat kita asumsikan tidak terjadi multikolinearitas pada model. Selanjutnya dengan menggunakan hipotesis uji, kita dapat memastikan apakah terjadi multikolinearitas atau tidak.

- Hipotesis Uji :  
 $H_0$  : tidak terdapat multikolinearitas antar variabel independen  
 $H_1$  : terdapat multikolinearitas antar variabel independen
- $\alpha: 5\% = 0.05$
- Statistik Uji :

```
> # Uji Multikolinearitas
> ols_vif_tol(modelSalaries_41)
Variables Tolerance VIF
1      V2 0.6518768 1.534032
2    dAsstP 0.5757156 1.736969
3    dProf 0.5415832 1.846438
4    dMale 0.9817707 1.018568
```

Nilai VIF tiap variabel bebas

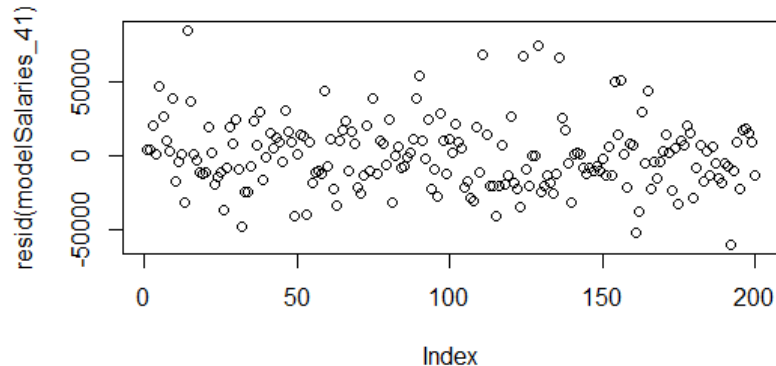
- Nilai VIF yrs.service : 1.534032
  - Nilai VIF dAsstP : 1.736969
  - Nilai VIF dPof : 1.846438
  - Nilai VIF dMale : 1.018568
- Daerah Kritis :  
 $H_0$  ditolak jika terdapat nilai  $VIF > 5$
  - Keputusan :

Berdasarkan hasil pengujian terlihat bahwa semua nilai  $VIF < 5$  maka  $H_0$  diterima. Ini berarti bahwa tidak terjadi multikolinearitas pada data.

### 9. Uji Autokorelasi

Uji ini digunakan untuk melihat apakah data observasi memiliki korelasi dengan data observasi sebelumnya.

- Melihat plot residual model



Secara visual, dapat dilihat bahwa tidak terdapat pola pada residual-residual data, sehingga dapat kita asumsikan tidak terjadi autokorelasi pada model. Selanjutnya dengan menggunakan hipotesis uji, kita dapat memastikan apakah terjadi autokorelasi atau tidak.

- Hipotesis Uji :  
 $H_0$  : tidak terdapat autokorelasi  
 $H_1$  : terdapat autokorelasi
- $\alpha: 5\% = 0.05$
- Statistik Uji :

```
> # Uji Autokorelasi
> dwtest(modelSalaries_41)

Durbin-watson test

data: modelSalaries_41
DW = 1.8469, p-value = 0.1365
alternative hypothesis: true autocorrelation is greater than 0
>
```

$Pvalue = 0.1365$

- Daerah Kritis :  
 $H_0$  ditolak jika  $Pvalue < 0.05$
- Keputusan :  
Berdasarkan hasil pengujian terlihat bahwa  $Pvalue = 0.1365 > 0.05$  maka  $H_0$  diterima. Ini berarti bahwa tidak terdapat autokorelasi pada model.

## 10. Perbaikan Model

Berdasarkan serangkaian pengujian di atas, model terbebas dari pelanggaran asumsi multikolinearitas dan autokorelasi, sedangkan untuk asumsi normalitas dan homoskedastisitas model mengalami pelanggaran sehingga diperlukan perbaikan model untuk mengatasi pelanggaran normalitas dan homoskedastisitas.

Langkah yang dilakukan untuk mengatasi pelanggaran normalitas pada kasus ini yaitu mentransformasikan data menjadi Moderate positive skewness dengan bentuk transformasi datanya adalah  $\text{SQRT}(x)$ . data yang ditransformasikan adalah variabel Salary menjadi "salary\_transformasi" dan yrs.service menjadi "yrs.service\_transformasi". Kemudian data yang telah ditransformasi dibentuk ke dalam model baru.

```
> # Mengatasi Pelanggaran Asumsi
>
> # Mengatasi Pelanggaran Normalitas
> salary_transformasi = sqrt(Salaries_41$V1)
> yrs.service_transformasi = sqrt(Salaries_41$V2)
> Salaries_sqrt <- cbind(Salaries_41$dAsstP,
  Salaries_41$dProf, Salaries_41$dMale, salary_transformasi, yrs.service_transformasi)
> Salaries_sqrt_41 <- data.frame(Salaries_sqrt)
> View(Salaries_sqrt_41)
>
> model_sqrt_41 = lm(salary_transformasi ~ ., data = Salaries_sqrt_41)
> summary(model_sqrt_41)

Call:
lm(formula = salary_transformasi ~ ., data = Salaries_sqrt_41)

Residuals:
    Min       1Q   Median       3Q      Max
-102.96  -22.59   -0.89    19.08   107.91

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    308.730     10.743   28.738 < 2e-16 ***
V1             -24.221      8.882   -2.727  0.00698 **
V2              52.796      6.971    7.573 1.42e-12 ***
V3              6.813      7.679    0.887  0.37608 .
yrs.service_transformasi -3.496      1.978   -1.767  0.07878 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 33.71 on 195 degrees of freedom
Multiple R-squared:  0.4156, Adjusted R-squared:  0.4036
F-statistic: 34.66 on 4 and 195 DF, p-value: < 2.2e-16
```

Selanjutnya akan dilakukan pengujian asumsi kembali model transformasi tersebut

```
>
> # Uji Normalitas
> resid_sqrt_41=resid(model_sqrt_41)
> shapiro.test(resid_sqrt_41) # berhasil normal

      Shapiro-wilk normality test

data:  resid_sqrt_41
W = 0.98632, p-value = 0.05028

>
>
> # Uji heterokedastisitas
> bptest(model_sqrt_41,studentize = F)

      Breusch-Pagan test

data:  model_sqrt_41
BP = 40.665, df = 4, p-value = 3.154e-08

>
> # Uji Multikolinearitas
> ols_vif_tol(model_sqrt_41)

```

	Variables	Tolerance	VIF
1	v1	0.5225633	1.913644
2	v2	0.5208540	1.919924
3	v3	0.9839563	1.016305
4	yrs.service_transformasi	0.5145080	1.943605

```
>
> # Uji Autokorelasi
> dwtest(model_sqrt_41)

      Durbin-watson test

data:  model_sqrt_41
DW = 1.859, p-value = 0.1566
alternative hypothesis: true autocorrelation is greater than 0

>
```

#### 10.1. Uji Normalitas

- Hipotesis Uji :

$H_0$  : residual berdistribusi normal

$H_1$  : residual tidak berdistribusi normal

- $\alpha$ : 5% = 0.05
- Statistik Uji :

$Pvalue = 0.05028$

- Daerah Kritis :

$H_0$  ditolak jika  $Pvalue < 0.05$

- Keputusan :

Berdasarkan hasil pengujian terlihat bahwa  $Pvalue = 0.05028 > 0.05$  maka  $H_0$  diterima. Ini berarti bahwa residual berdistribusi normal. Sehingga pelanggaran asumsi normalitas berhasil teratasi.

#### 10.2. Uji Heteroskedastisitas

- Hipotesis Uji :

$$H_0 : \sigma_i^2 = \sigma^2 \quad i = 0,1,2,3,4 \text{ (tidak terjadi heteroskedastisitas)}$$

$$H_1 : \exists \sigma_i^2 \neq \sigma^2 \quad i = 0,1,2,3,4 \text{ (terjadi heteroskedastisitas)}$$

- $\alpha: 5\% = 0.05$
- Statistik Uji :

$$Pvalue = 3.154e - 08$$

- Daerah Kritis :

$$H_0 \text{ ditolak jika } Pvalue < 0.05$$

- Keputusan :

Berdasarkan hasil pengujian terlihat bahwa  $Pvalue = 3.154e - 08 < 0.05$  maka  $H_0$  ditolak. Ini berarti bahwa masih terjadi heteroskedastisitas pada model transformasi.

- Namun berdasarkan output summary model transformasi terlihat bahwa nilai residual std error nya berubah, yang semula bernilai 23610 menjadi 33.71 . Hal ini menunjukkan bahwa nilai prediksi yang dihasilkan oleh model transformasi jauh lebih dekat dengan pengamatan sebenarnya dibandingkan dengan nilai prediksi yang dihasilkan oleh model awal. Sehingga dapat dikatakan bahwa pelanggaran homoskedastisitas juga telah teratasi pada model transformasi ini.

#### 10.3. Uji Multikolinearitas

- Hipotesis Uji :

$$H_0 : \text{tidak terdapat multikolinearitas antar variabel independen}$$

$$H_1 : \text{terdapat multikolinearitas antar variabel independen}$$

- $\alpha: 5\% = 0.05$
- Statistik Uji :

Nilai VIF tiap variabel bebas

- Nilai VIF yrs.service\_transformasi : 1.943605
- Nilai VIF dAsstP : 1.913644
- Nilai VIF dPof : 1.919924
- Nilai VIF dMale : 1.016305

- Daerah Kritis :

$$H_0 \text{ ditolak jika terdapat nilai } VIF > 5$$

- Keputusan :  
Berdasarkan hasil pengujian terlihat bahwa semua nilai  $VIF < 5$  maka  $H_0$  diterima. Ini berarti bahwa tidak terjadi multikolinearitas pada data transformasi.
- 10.4. Uji Autokorelasi
- Hipotesis Uji :  
 $H_0$  : tidak terdapat autokorelasi  
 $H_1$  : terdapat autokorelasi
  - $\alpha: 5\% = 0.05$
  - Statistik Uji :  
 $Pvalue = 0.1566$
  - Daerah Kritis :  
 $H_0$  ditolak jika  $Pvalue < 0.05$
  - Keputusan :  
Berdasarkan hasil pengujian terlihat bahwa  $Pvalue = 0.1365 > 0.05$  maka  $H_0$  diterima. Ini berarti bahwa tidak terdapat autokorelasi pada model transformasi.

## 11. Membentuk Model Final

Setelah melakukan pengujian asumsi ulang pada model transformasi, semua asumsi baik normalitas, homoskedastisitas, multikolinearitas, hingga autokorelasi, diperoleh hasil bahwa sudah tidak terdapat pelanggaran. Sehingga dapat dikatakan bahwa model sudah layak untuk digunakan secara praktis, dengan model final sebagai berikut :

```
> model_sqrt_41 = lm(salary_transformasi ~ ., data = Salaries_sqrt_41)
> summary(model_sqrt_41)

Call:
lm(formula = salary_transformasi ~ ., data = Salaries_sqrt_41)

Residuals:
    Min       1Q   Median       3Q      Max
-102.96  -22.59   -0.89   19.08   107.91

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    308.730     10.743   28.738 < 2e-16 ***
v1             -24.221      8.882   -2.727  0.00698 **
v2              52.796      6.971    7.573 1.42e-12 ***
v3              6.813      7.679    0.887  0.37608
yrs.service_transformasi -3.496      1.978   -1.767  0.07878 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 33.71 on 195 degrees of freedom
Multiple R-squared:  0.4156, Adjusted R-squared:  0.4036
F-statistic: 34.66 on 4 and 195 DF, p-value: < 2.2e-16
```

### Model Regresi Final :

- Model regresi keseluruhan :  $\text{salary\_transformasi} = 308.730 - 24.221 V1 + 52.796 V2 + 6.813 V3 - 3.496 \text{ yrs.service\_transformasi} + e$   
Dengan  $\text{salary\_transformasi} = \text{salary}$ ,  $V1 = \text{assistant professor}$ ,  $V2 = \text{professor}$ ,  $V3 = \text{laki-laki}$ ,  $\text{yrs.service\_transformasi} = \text{lama kerja}$
- Model untuk rank Professor laki-laki ( $\text{Prof}=1, \text{AsstProf}=0, \text{Male}=1, \text{Female}=0$ )  
 $\text{salary\_transformasi} = 308.730 - 3.496 \text{ yrs.service\_transformasi} + 52.796 + 6.813 + e$   
 $\text{salary\_transformasi} = 368.330 - 3.496 \text{ yrs.service\_transformasi} + e$
- Model untuk rank Professor perempuan ( $\text{Prof}=1, \text{AsstProf}=0, \text{Male}=1, \text{Female}=0$ )  
 $\text{salary\_transformasi} = 308.730 - 3.496 \text{ yrs.service\_transformasi} + 52.796 + e$   
 $\text{salary\_transformasi} = 361.526 - 3.496 \text{ yrs.service\_transformasi} + e$
- Model untuk rank Assistant Professor laki-laki ( $\text{Prof}=0, \text{AsstProf}=1, \text{Male}=1, \text{Female}=0$ )  
 $\text{salary\_transformasi} = 308.730 - 3.496 \text{ yrs.service\_transformasi} - 24.221 + 6.813 + e$   
 $\text{salary\_transformasi} = 291.322 - 3.496 \text{ yrs.service\_transformasi} + e$
- Model untuk rank Assistant Professor perempuan ( $\text{Prof}=0, \text{AsstProf}=1, \text{Male}=1, \text{Female}=0$ )  
 $\text{salary\_transformasi} = 308.730 - 3.496 \text{ yrs.service\_transformasi} - 24.221 + e$   
 $\text{salary\_transformasi} = 284.509 - 3.496 \text{ yrs.service\_transformasi} + e$
- Model untuk rank Associate Professor laki-laki ( $\text{Prof}=1, \text{AsstProf}=1, \text{Male}=1, \text{Female}=0$ )  
 $\text{salary\_transformasi} = 308.730 - 3.496 \text{ yrs.service\_transformasi} + 6.813 + e$   
 $\text{salary\_transformasi} = 315.543 - 3.496 \text{ yrs.service\_transformasi} + e$
- Model untuk rank Associate Professor perempuan ( $\text{Prof}=1, \text{AsstProf}=1, \text{Male}=1, \text{Female}=0$ )  
 $\text{salary\_transformasi} = 308.730 - 3.496 \text{ yrs.service\_transformasi} + e$

### Interpretasi :

- Berdasarkan output yang dihasilkan didapat nilai koefisien  $b_0, b_1, b_2, b_3, b_4$  berturut-turut adalah 308.730, -24.221, 52.796, 6.813, -3.496
- Nilai dugaan intercept = 308.730, artinya rata-rata gaji associate professor perempuan sebesar 308.730 ketika lama kerja=0.
- Nilai dugaan  $\text{yrs.service} = -3.496$  artinya rata-rata perubahan gaji tiap staff sebesar 3.496 ketika terjadi perubahan 1 lama kerja.
- Nilai dugaan  $d\text{AsstP} = -24.221$ , artinya rata-rata perbedaan gaji assistant professor dengan associate professor untuk gender yang sama adalah 24.221 ketika lama kerja=0.
- Nilai dugaan  $d\text{Prof} = 52.796$ , artinya rata-rata perbedaan gaji professor dengan associate professor untuk gender yang sama adalah 52.796 ketika lama kerja=0.
- Nilai dugaan  $d\text{Male} = 6.813$ , artinya rata-rata perbedaan gaji staff laki-laki dan perempuan untuk level yang sebanding adalah 6.813 ketika lama kerja=0.
- Adjusted R-squared : 0.4036, artinya sebesar 40.36% model mampu menjelaskan keragaman variabel salary dan sekitar 59.64% keragaman yang tidak mampu dijelaskan oleh variabel penjelas dan terletak pada komponen error.

## 12. KESIMPULAN

- Dari 200 staff yang terpilih dalam data sample, persebaran level dan gender nya yaitu sebanyak 30 orang Associate Prof laki-laki, 5 orang Associate Prof perempuan, 27 orang Assistant Prof laki-laki, 6 orang Assistant Prof perempuan, 121 orang professor laki-laki, dan 11 orang professor perempuan.
- Untuk k-1 peubah dummy, peubah rank dibentuk menjadi 2 peubah dummy dengan referensi AssocPorf=0 yaitu dAsstP=1 dan dProf=1, dan peubah sex dibentuk menjadi 1 peubah dummy dengan referensi Female=0 yaitu dMale=1.
- Secara keseluruhan model regresi awal yang digunakan adalah  $\text{Salary} = 91369.0 - 258.8 \text{ lama kerja} - 13343.0 \text{ Assistant Professor} + 35190.9 \text{ Professor} + 4946.8 \text{ Male} + e$
- Berdasarkan serangkaian pengujian asumsi klasik, model terbebas dari pelanggaran asumsi multikolinearitas dan autokorelasi, sedangkan untuk asumsi normalitas dan homoskedastisitas model mengalami pelanggaran.
- Setelah melakukan pengujian asumsi ulang pada model transformasi dan semua pelanggaran telah teratasi, diperoleh model final yaitu  $\text{Salary} = 308.730 - 24.221 \text{ Assistant Professor} + 52.796 \text{ Professor} + 6.813 \text{ laki-laki} - 3.496 \text{ lama kerja} + e$ .
- Dengan  $d_{\text{justed R-squared}} : 0.4036$ , artinya sebesar 40.36% model mampu menjelaskan keragaman variabel salary dan sekitar 59.64% keragaman yang tidak mampu dijelaskan oleh variabel penjelas dan terletak pada komponen error.