LAPORAN TUGAS AKHIR PROYEK METODE STATISTIKA ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI HARGA TIKET KEBERANGKATAN PULANG KAMPUNG MAHASISWA TEKNOLOGI SAINS DATA ANGKATAN 2020, 2021, 2022 UNIVERSITAS AIRLANGGA



Kelas:

SD - A2

Dosen Pengampu:

Ratih Ardiati Ningrum, S.Si., M.Stat.

Disusun Oleh:

| Manda Diana Putri | 164221048 |
|-------------------------|-----------|
| Nismara Andini | 164221075 |
| Fadli Muhammad | 164221081 |
| Hadya Adira Perdana | 164221085 |
| Ananda Alvin Bernerdian | 164221096 |

UNIVERSITAS AIRLANGGA SURABAYA 2023

DAFTAR ISI

| DAFTAR ISI | II |
|--------------------------------------|-----|
| DAFTAR TABEL | III |
| DAFTAR GAMBAR | III |
| BAB 1 | 1 |
| PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 1 |
| 1.3 Tujuan | 1 |
| 1.4 Batasan Penelitian | 2 |
| BAB II | 3 |
| Tinjauan Pustaka | 3 |
| 2.1 Tinjauan Statistika | 3 |
| 2.2 Tinjauan Non Statistika | 4 |
| BAB III | 6 |
| METODOLOGI | 6 |
| 3.1 Metode Pengambilan Sampel | 6 |
| 3.2 Variabel Penelitian | 6 |
| BAB IV | 7 |
| HASIL DAN ANALISIS | 7 |
| 4.1 Analisis Eksplorasi Data | 7 |
| 4.2 Analisis Regresi Linear Berganda | 8 |
| BAB V | 14 |
| KESIMPULAN | 14 |
| 5.1 Kesimpulan | 14 |
| 5.2 Saran | 15 |
| DAFTAR PUSTAKA | 16 |
| LAMPIRAN | 17 |

DAFTAR TABEL

| BAB 4 | |
|---|------------|
| 4.1 Korelasi Antar Variabel Correlations | 8 |
| 4.2 Regresi Variabel Dependen dan Independen | |
| Model Summary b | 8 |
| 4.3 Regresi Variabel Dependen dan Independen | |
| Model Unstandardized B | 9 |
| 4.4 Uji Serentak ANOVA a | 9 |
| 4.5 Uji Parsial Coefficients a | 10 |
| 4.6 Uji Multikolinearitas Coefficients a | 10 |
| 4.7 Uji Normalitas One-Sample Kolmogorov-Smirnov Test | 6 1 |
| 4.8 Uji Autokorelasi Model Summary b | 6 2 |
| 4.9 Model Regresi Variabel Selection Coefficients a | 13 |
| 4.10 Kebaikan Model Regresi Akhir Model Summary f | 13 |
| DAFTAR GAMBAR | |
| BAB 4 | |
| 4.1 Visualisasi Hubungan antar variabel | 7 |
| 4.2 Boxplot Harga Tiket Pulang | 7 |
| 4.3 Scatter Plot seluruh Variabel | 7 |
| 4.4 Scatter Plot Uji Normalitas | 11 |
| 4.5 Scatter Plot Uji Heteroskedastisitas | 12 |

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Dalam dunia perkuliahan terdapat perbedaan yang paling besar dalam lingkup sosial, dimana banyak pendatang dari luar kota yang memiliki tujuan yang sama yaitu mengenyam pendidikan. Secara tidak langsung ketika memiliki libur yang cukup panjang, maka pulang kampung sudah menjadi hal umum yang dilakukan oleh mahasiswa. Dorongan untuk pulang kampung ini tidak lain adalah mengunjungi orang tua dan kerabat terdekat. Untuk pergi ke kampung halaman, merantau membutuhkan moda transportasi terbaik untuk diri mereka sendiri, pemilihan moda ini didasarkan terhadap kemampuan diri masing-masing.

Harga tiket merupakan salah satu faktor penting dalam pemilihan moda transportasi. Finansial setiap pribadi berbeda-beda sedangkan harga tiket berubah-ubah yang menyebabkan sulitnya pulang dengan terencana. Peningkatan harga tiket yang tidak terduga dapat menyebabkan ketidakpastian dan kesulitan bagi mahasiswa dalam mempersiapkan perjalanan pulang kampung mereka. Ketidakmampuan untuk merencanakan perjalanan dengan baik dapat berdampak negatif pada kesejahteraan mahasiswa dan hubungan keluarga, serta mengganggu keseimbangan keuangan pribadi.

Dalam hal ini penting untuk mengetahui faktor-faktor apa saja yang mempengaruhi harga tiket transportasi untuk pulang kampung. Dengan memahami faktor-faktor ini, akan memungkinkan untuk mengidentifikasi penyebab kenaikan harga tiket dan mencari solusi yang tepat guna mengurangi dampaknya.

1.2 Rumusan Masalah

- 1. Bagaimana hubungan antara harga tiket sebagai variabel dependen terhadap variabel independen?
- 2. Apakah variabel prediktor memiliki pengaruh terhadap harga tiket?
- 3. Bagaimana hasil analisis regresi antara faktor-faktor yang mempengaruhi harga tiket dan harga tiket itu sendiri?

1.3 Tujuan

- 1. Mengetahui hubungan antara harga tiket terhadap variabel independen.
- 2. Mengetahui faktor-faktor yang mempengaruhi harga tiket memiliki pengaruh terhadap harga tersebut.
- 3. Mengetahui hasil analisis regresi antara faktor-faktor yang mempengaruhi harga tiket dan harga tiket itu sendiri.

1.4 Batasan Penelitian

Batasan penelitian dilakukan untuk memastikan fokus dan ruang lingkup penelitian yang jelas serta tidak meluas ke pembahasan yang lain. Oleh karena itu, beberapa aspek yang menjadi batasan penelitian ini antara lain:

A. Responden

Penelitian ini akan difokuskan pada populasi mahasiswa Teknologi Sains Data Universitas Airlangga sebagai sampel yang mewakili mahasiswa pendatang dari luar kota yang melakukan perjalanan pulang kampung. Dan kami mendapatkan 80 orang sebagai sampel dengan menggunakan rumus *Slovin*, Rumus *Slovin*.

B. Jenis transportasi

Penelitian ini akan mempertimbangkan empat jenis transportasi utama yang sering digunakan oleh mahasiswa dalam perjalanan pulang kampung, yaitu pesawat, kapal, kereta api, dan bus. Faktor-faktor yang mempengaruhi harga tiket akan dianalisis berdasarkan keempat jenis transportasi ini.

C. Variabel yang akan diteliti

Variabel yang akan diteliti terbatas pada faktor-faktor yang mempengaruhi harga tiket, seperti jenis transportasi, jarak ke tempat tujuan, waktu tempuh, musim keberangkatan, kelas transportasi, dan waktu pembelian tiket. Faktor-faktor lain yang tidak langsung berhubungan dengan harga tiket akan diabaikan dalam penelitian ini.

Dengan membatasi aspek-aspek tersebut, penelitian dapat lebih terfokus pada variabel yang relevan dan memungkinkan analisis yang lebih mendalam mengenai faktor-faktor yang mempengaruhi harga tiket keberangkatan pulang kampung mahasiswa Teknologi Sains Data Universitas Airlangga.

BABII

Tinjauan Pustaka

2.1 Tinjauan Statistika

2.1.1 Penerapan statistika

Penggunaan statistika dalam mengolah data penelitian berpengaruh terhadap tingkat analisis hasil penelitian. Penelitian-penelitian dalam yang menggunakan perhitungan-perhitungan statistika, akan menghasilkan data yang mendekati benar jika memperhatikan tata cara analisis data yang digunakan. Dalam memprediksi dan mengukur nilai dari pengaruh satu variabel (bebas/independen/prediktor) terhadap variabel lain (tak bebas/dependent/response) dapat digunakan uji regresi.

2.1.2 Variabel

2.1.2.1 Variabel Dependen

Variabel dependen adalah variabel yang mengandung perubahan yang menjadi tujuan pengamatan. Variabel ini disebut juga sebagai variabel respon. Umumnya, setiap variabel memiliki nilai yang berbeda-beda dan terukur.

2.1.2.2 Variabel Independen

Variabel independen adalah variabel yang mempengaruhi atau menjadi penyebab berubahnya variabel dependen. Variabel ini mempengaruhi faktor-faktor yang diukur atau diamati seorang peneliti untuk menentukan hubungan antara fenomena yang sedang diamati

2.1.2.3 Variabel Dummy

Variabel dummy merupakan salah satu variabel yang digunakan untuk membuat data yang bersifat kualitatif. Analisis regresi berganda dengan variabel dummy merupakan salah satu metode pemodelan yang sederhana untuk menentukan sebab akibat antara variabel dependen dengan variabel independen. Analisis ini bertujuan menghasilkan model untuk melakukan prediksi dan mendeteksi variabel independen (Y) yang memberikan pengaruh terhadap variabel dependen (X). (Novrika, 2020)

2.1.2 Uji regresi

Uji regresi merupakan hubungan antara satu variabel dengan satu atau lebih variabel lain. Apabila variabel bebas hanya satu, maka uji regresinya disebut dengan uji regresi sederhana. Apabila variabel bebas lebih dari pada satu, maka uji regresi disebut dengan uji regresi linier berganda. Dikatakan linier berganda karena terdapat dua atau lebih variabel bebas yang mempengaruhi variabel tak bebas.

2.1.3 Regresi Linier Berganda

Regresi linier berganda merupakan model persamaan yang menjelaskan hubungan satu variabel dependen (Y) dengan dua atau lebih variabel dependen (X1, X2, ..., Xn). Tujuan dari uji regresi linier berganda adalah untuk memprediksi nilai variabel respon (Y) apabila nilai-nilai variabel bebasnya/ predictor (X1, X2, ..., Xn) diketahui. Disamping itu juga untuk dapat mengetahui bagaimanakah arah hubungan variabel respon dengan variabel prediktor.

Persamaan regresi linier berganda dituliskan:

$$Y = B0 + B1X1 + B2X2 + ... + Bnxn + e$$

keterangan:

Y = Variabel dependen / Variabel respon

B0 = Konstanta

Bn = Slope / Koefisien estimasi

Xn = Variabel independen / Variabel prediktor

e = error term

2.1.4 Uji signifikansi

Uji signifikan parameter ini dilakukan untuk mengetahui apakah taksiran parameter berpengaruh terhadap model atau tidak secara signifikan, serta mengetahui seberapa besar pengaruh masing-masing parameter tersebut. Cara pengujian signifikansi parameter terbagi menjadi dua, yaitu secara serentak dan secara parsial.

2.1.5 Uji Parsial

Uji parsial atau individual adalah untuk menguji apakah suatu variabel bebas berpengaruh atau tidak terhadap variabel tidak bebas. Tujuan pengujian ini adalah untuk mengetahui apakah masing-masing variabel independen mempengaruhi variabel dependen secara signifikan.

2.1.6 Uji Serentak

Uji serentak dilakukan untuk mengetahui signifikansi parameter model regresi secara bersama-sama menggunakan analisis ANOVA (Analysis of Variance).

2.1.7 OLS

Ordinary Least Squares (OLS) adalah sebuah teknik regresi yang termasuk ke dalam *supervised learning*. Ini adalah *method* dari python yang digunakan untuk mengestimasikan sebuah parameter yang tidak diketahui dengan membuat model yang akan meminimalkan *sum of the squared errors* antara variabel yang diamati atau variabel respons dengan variabel yang diprediksi.

2.2 Tinjauan Non Statistika

2.2.1 Tiket

Tiket didefinisikan sebagai dokumen kecil dalam bentuk cetak maupun digital sebagai tanda telah membayar ongkos dan sebagainya. Tiket

digunakan sebagai bukti bahwa seseorang telah membayar suatu jasa untuk ditukarkan ke pemilik jasa.

2.2.2 Transportasi

Menurut Utomo (2010:25-35), transportasi adalah: (1) pemindahan barang dan manusia dari tempat asal ke tempat tujuan, (2) salah satu jenis kegiatan yang menyangkut peningkatan kebutuhan manusia dengan mengubah letak geografis barang dan orang sehingga akan menimbulkan adanya transaksi. Sedangkan menurut Sukarto (2006; 94) transportasi adalah perpindahan dari suatu tempat ke tempat lain dengan menggunakan alat pengangkutan, baik yang digerakkan oleh tenaga manusia, hewan (kuda, sapi, kerbau), atau mesin. Konsep transportasi didasarkan pada adanya perjalanan (trip) antara asal (origin) dan tujuan (destination). Transportasi umum yang kerap digunakan oleh mahasiswa TSD untuk pulang kampung antara lain Pesawat Udara, Kereta Api, Kapal, dan Bis. Pesawat Udara Merupakan alat yang dapat terbang di atmosfer karena daya angkat dari reaksi udara. Kereta Api adalah sarana transportasi berupa kendaraan dengan tenaga gerak, baik berjalan sendiri maupun dirangkaikan dengan kendaraan lainnya, yang akan ataupun sedang bergerak di rel. Kapal, adalah kendaraan pengangkut penumpang dan barang di laut (sungai dsb). Bis adalah Kendaraan berbentuk kotak yang berkapasitas puluhan orang yang memiliki fungsi untuk memindahkan penumpang dari satu tempat ke tempat lainnya. Transportasi umum tersebut memiliki kelas berbeda yang ditawarkan berdasarkan tingkat kenyamanan yang berbeda antar kelas nya.

2.2.3 Jarak

Jarak adalah suatu ukuran numeric yang menunjukkan seberapa jauh posisi suatu objek dengan objek lainnya.

2.2.4 Waktu

Waktu adalah serangkaian saat ketika proses suatu kejadian, perubahan, atau keadaan saat berlangsung suatu benda, lamanya saat tertentu untuk melakukan suatu, sebuah kesempatan tempo, peluang, ketika saat, keadaan hari dan saat yang ditentukan berdasarkan pembagian benda dunia.

BAB III METODOLOGI

3.1 Metode Pengambilan Sampel

3.1.1 Populasi

Populasi yang kami gunakan pada analisis ini adalah mahasiswa Teknologi Sains Data Universitas Airlangga angkatan 2020, 2021, 2022 dengan total mahasiswa 318 orang.

3.1.2 Metode Sampling yang Digunakan

Metode pengambilan sampel yang kami gunakan menggunakan metode slovin dengan teknik simple random sampling (SRS). Metode Slovin merupakan metode untuk sampel yang representatif dari populasi yang lebih besar Alasan kami memilih metode ini karena untuk sampel minimumnya ditentukan sesuai yang diperlukan dalam penelitian dan sampel yang didapat sedikit tetapi dapat mewakili keseluruhan populasi. Berikut rumus slovin yang kami gunakan untuk menentukan ukuran sampel:

$$n = \frac{N}{1 + N(e^2)}$$

$$n = \frac{318}{1 + 318(0,1^2)} = 76,076555 \sim 77$$

Keterangan:

n = jumlah sampel yang dicari

N = jumlah populasi (mahasiswa TSD Universitas Airlangga angkatan 2020-2022 = 318)

e = margin of error (10% = 0.1)

3.1.3 Tingkat Signifikansi

Pada penelitian ini kami menggunakan tingkat signifikansi sebesar 10% agar ukuran sampel yang digunakan tidak terlalu banyak, melihat estimasi populasi mencapai 318 mahasiswa.

3.2 Variabel Penelitian

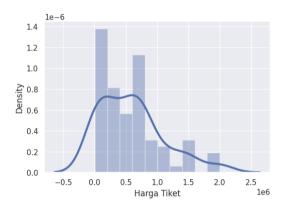
Pada penelitian ini kami menggunakan dua variabel, yaitu variabel independen sebagai prediktor dan variabel dependen sebagai respon. Variabel dependen yang kami gunakan adalah harga tiket pulang dan variabel independen ada 6 variabel, yakni:

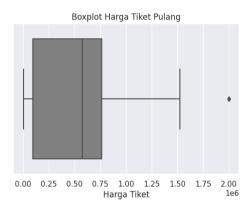
- 1. Transportasi yang dipilih untuk pulang
- 2. Jarak dari tempat tinggal di Surabaya ke tempat tujuan (dalam Km)
- 3. Waktu tempuh perjalanan (dalam Jam)
- 4. Musim keberangkatan
- 5. Kelas transportasi yang digunakan
- 6. Waktu Pembelian Tiket

BAB IV HASIL DAN ANALISIS

4.1 Analisis Eksplorasi Data

4.1.1 Menyajikan secara visual hubungan antar variabel



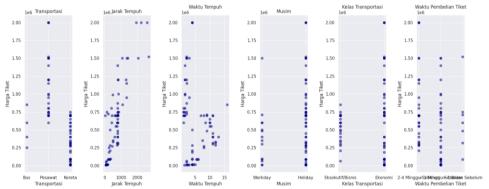


(Gambar 4.1: Visualisasi Histogram)

(Gambar 4.2: Boxplot)

Dari hasil grafik Gambar 4.1, kita dapatkan bahwa dari 80 sampel menunjukkan grafik tersebut bersifat right-skewness, Hal ini menunjukkan adanya data yang lebih ekstrim atau langka di sisi kanan distribusi yaitu ada pada data rentang 1.500.000 - 2.000.000 Rupiah. Sedangkan frekuensi data terbanyak terdapat pada rentang 0 - 1.000.000 Rupiah.

Dari Boxplot pada Gambar 4.2 ada 25% data harga tiket ada di bawah 100.000, 50% data harga tiket berada di bawah 600.000, dan 75% data harga tiket ada di bawah 750.00. Serta terdapat outlier data yang berada di harga 2.000.000 karena adanya outlier diduga memiliki distribusi yang kurang normal.



(Gambar 4.3: Scatter Plot seluruh Variabel)

4.1.2 Menghitung korelasi antar variabel

Correlations

| | | Hargatiketpul angketempatt inggalcontoh 750000 | Jarakdaritem pattinggaldiS urabayakete mpattujuanda lamKm | Waktutempuh perjalanandal amJam | Transportasiy angdipilihunt ukpulang=Bu s | Transportasiy angdipilihunt ukpulang=Ker eta | Transportasiy angdipilihunt ukpulang=Pe sawat |
|---------------------|------------------------|---|---|---------------------------------------|--|---|--|
| Pearson Correlation | Hargatiketpulangketemp | 1.000 | .856 | 207 | 030 | 731 | .765 |

| Musimkebera ngkatan=Holi day | Musimkebera ngkatan=Wor kday | Kelastranspo rtasiyangdigu nakan=Ekono mi | Kelastranspo rtasiyangdigu nakan=Eksek utif/Bisnis | WaktuPembel ianTiket=2-4 Minggu Sebelum | WaktuPembel ianTiket=< 1 Minggu Sebelum | WaktuPembel ianTiket=> 1 Bulan Sebelum |
|------------------------------------|------------------------------------|--|---|--|--|---|
| .166 | 166 | .079 | 079 | .366 | 373 | .028 |

(Tabel 4.1: Korelasi Antar Variabel Correlations)

Angka signifikansi sebesar 0,1 mempunyai pengertian bahwa kami menggunakan tingkat kepercayaan 90%. Interpretasi signifikansi analisis korelasi akan sangat tergantung pada tingkat kepercayaan yang telah kita pilih atau tentukan di awal.

Hubungan/ korelasi harga tiket pulang ke tempat tinggal sebagai variabel dependen dengan variabel independen menunjukkan bahwa:

- a. Hubungan yang kuat dan positif / searah (nilai korelasi > 0,5) adalah jarak
- b. Hubungan yang lemah dan negatif / berlawanan arah (-0,5 < nilai korelasi < 0) adalah waktu tempuh, transportasi bus, transportasi kereta, transportasi pesawat, musim keberangkatan workday, kelas transportasi bisnis, dan pembelian tiket =< 1 minggu.
- c. Hubungan yang lemah dan positif / searah (0<nilai korelasi <0,5) adalah musim keberangkatan holiday, kelas transportasi ekonomi, pembelian tiket 2-4 minggu, dan pembelian tiket => 1 bulan

4.2 Analisis Regresi Linear Berganda

4.2.1 Regresi Variabel Dependen dan Independen

Model Summary^b

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin- Watson |
|-------|-------------------|----------|----------------------|-------------------------------|-------------------|
| 1 | .949 ^a | .900 | .889 | 170970.614 | 1.995 |

- a. Predictors: (Constant), Waktu_Pembelian_Tiket=> 1 Bulan Sebelum, Transportasi=Pesawat, Transportasi=Bus, Musim=Holiday, Waktu_Pembelian_Tiket=2-4 Minggu Sebelum, Kelas_Transportasi=Eksekutif/Bisnis, Jarak_Tempuh, Waktu_Tempuh
- b. Dependent Variable: Harga_Tiket

(Tabel 4.2: Regresi Variabel Dependen dan Independen Model Summary b)

Nilai R-Squared yang didapatkan adalah 0,90. Artinya, model mampu menjelaskan variansi variabel respon sebesar 90%, sedangkan sisanya yaitu 10% dijelaskan oleh variabel lain yang tidak diteliti. Secara keseluruhan, model ini memiliki tingkat kebaikan yang tinggi karena

mampu menjelaskan sebagian besar variabilitas variabel respon yang diamati.

| Model | | Unstandardized B |
|-------|--|---------------------|
| 1 | (Constant) | -77024.221 |
| | Jarak_Tempuh | 551.920 |
| | Waktu_Tempuh | 2915.193 |
| | Transportasi=Bus | -10015.639 |
| | Transportasi=Pesawat | 614216.649 |
| | Musim=Holiday | 47031.558 |
| | Kelas_Transportasi=Eks ekutif/Bisnis | 273509.379 |
| | Waktu_Pembelian_Tiket= 2-4 Minggu Sebelum | -87809.598 |
| | Waktu_Pembelian_Tiket= > 1 Bulan Sebelum | -177877.068 |

(Tabel 4.3: Regresi Variabel Dependen dan Independen

Model Unstandardized B)

Kami melakukan analisis linier berganda pada SPSS didapatkan hasil semua persamaan seperti gambar diatas. Dari tabel tersebut persamaan regresi yang diperoleh, yakni:

```
Y = -77024.221 + 551.920 (hargatiket pulangketempattinggal) + 2915.193 (Waktu Tempuh) - 100015.639 (Transportasi: Bus) + 614216.649 (Transportasi: Pesawat) + 47031.558 (Musim: Holiday) + 273509.379 (Kelas Transportasi: Eksekutif / Bisnis) - 87809.598 (Waktu Pembelian Tiket: <math>2 - 4 Minggu Sebelum) - 177877.068 (Waktu Pembelian Tiket: > 1 Bulan Sebelum)
```

4.2.2 Uji Signifikansi Parameter

4.2.2.1 Uji Serentak

ANOVA^a

| del | | Sum of Squares | df | Mean Square | F | Sig. |
|-----|------------|-------------------|----|-------------|--------|-------------------|
| | Regression | 1.872E+13 | 8 | 2.340E+12 | 80.039 | .000 ^b |
| | Residual | 2.075E+12 | 71 | 2.923E+10 | | |
| | Total | 2.079E+13 | 79 | | | |

a. Dependent Variable: Harga_Tiket

(Tabel 4.4: Uji Serentak ANOVA)

Uji serentak merupakan pengujian untuk melihat apakah variabel independen secara bersama-sama serentak mempunyai pengaruh yang signifikan terhadap variabel dependen.

Dari tabel di atas P-Value memiliki nilai 0.000. Dapat diambil kesimpulan P-Value 0.000<Alpha 0.1. Artinya, terdapat hubungan antara variabel independen yang diteliti dengan variabel dependen(Harga Tiket Pulang Tempat Tinggal)

b. Predictors: (Constant), Waktu_Pembelian_Tiket=> 1 Bulan Sebelum, Transportasi=Pesawat, Transportasi=Bus, Musim=Holiday, Waktu_Pembelian_Tiket=2-4 Minggu Sebelum, Kelas_Transportasi=Eksekutif/Bisnis, Jarak_Tempuh, Waktu_Tempuh

4.2.2.2 Uji Parsial

| | | | | Co | efficients ^a |
|--|---------------|----------------|------------------------------|--------|-------------------------|
| | Unstandardize | d Coefficients | Standardized Coefficients | | |
| | В | Std. Error | Beta | t | Sig. |
| (Constant) | -77024.221 | 66960.748 | | -1.150 | .254 |
| Jarak_Tempuh | 551.920 | 53.147 | .598 | 10.385 | .000 |
| Waktu_Tempuh | 2915.193 | 8861.605 | .022 | .329 | .743 |
| Transportasi=Bus | -10015.639 | 95535.820 | 004 | 105 | .917 |
| Transportasi=Pesawat | 614216.649 | 80272.421 | .575 | 7.652 | .000 |
| Musim=Holiday | 47031.558 | 54350.836 | .035 | .865 | .390 |
| Kelas_Transportasi=Eks ekutif/Bisnis | 273509.379 | 58148.033 | .246 | 4.704 | .000 |
| Waktu_Pembelian_Tiket= 2-4 Minggu Sebelum | -87809.598 | 54367.949 | 084 | -1.615 | .111 |
| Waktu_Pembelian_Tiket= > 1 Bulan Sebelum | -177877.068 | 77212.567 | 099 | -2.304 | .024 |

(Tabel 4.5: Uji Parsial Coefficients)

Berdasarkan tabel diatas variabel Jarak Tempuh, Transportasi=Pesawat, Kelas Transportasi=Eksekutif/Bisnis, Waktu Pembelian Tiket=>1 Bulan Sebelum memiliki P-Value < Alpha 0,1. Maka Tolak H0. Artinya, terdapat hubungan signifikan antara variabel-variabel tersebut dengan variabel dependen yaitu Harga Tiket Pulang Tempat Tinggal

Sedangkan variabel Waktu Tempuh, Transportasi=Bus, Musim=Holiday, Waktu Pembelian Tiket=2-4 Minggu Sebelum. Didapatkan P-Value > Alpha 0,1. Maka Gagal Tolak H0. Artinya, tidak terdapat hubungan signifikan antara variabel-variabel tersebut dengan variabel dependen yaitu Harga Tiket Pulang Tempat Tinggal.

4.2.3 Uji Asumsi Error

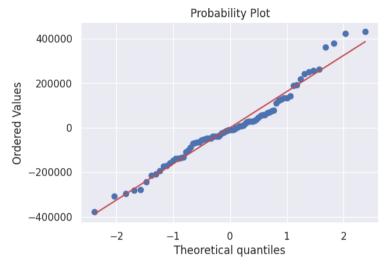
4.2.3.1 Uji Multikolinearita

| | Coefficients ^a | | | | | | | | | | | | | |
|--|---------------------------|----------------|------------------------------|--------|------|---------------------------------|-------------|--------------|---------|------|-------------------------|------|--|--|
| | Unstandardize | d Coefficients | Standardized Coefficients | | | 90.0% Confidence Interval for B | | Correlations | | | Collinearity Statistics | | | |
| | В | Std. Error | Beta | t | Sig. | Lower Bound | Upper Bound | Zero-order | Partial | Part | Tolerance | VIF | | |
| (Constant) | -77024.221 | 66960.748 | | -1.150 | .254 | -188620.981 | 34572.538 | | | | | | | |
| Jarak_Tempuh | 551.920 | 53.147 | .598 | 10.385 | .000 | 463.345 | 640.496 | .856 | .777 | .389 | .424 | 2.36 | | |
| Waktu_Tempuh | 2915.193 | 8861.605 | .022 | .329 | .743 | -11853.556 | 17683.941 | 207 | .039 | .012 | .313 | 3.19 | | |
| Transportasi=Bus | -10015.639 | 95535.820 | 004 | 105 | .917 | -169235.603 | 149204.326 | 030 | 012 | 004 | .843 | 1.18 | | |
| Transportasi=Pesawat | 614216.649 | 80272.421 | .575 | 7.652 | .000 | 480434.660 | 747998.639 | .765 | .672 | .287 | .249 | 4.01 | | |
| Musim=Holiday | 47031.558 | 54350.836 | .035 | .865 | .390 | -43549.527 | 137612.643 | .166 | .102 | .032 | .857 | 1.16 | | |
| Kelas_Transportasi=Eks ekutif/Bisnis | 273509.379 | 58148.033 | .246 | 4.704 | .000 | 176599.887 | 370418.872 | 079 | .487 | .176 | .515 | 1.94 | | |
| Waktu_Pembelian_Tiket= 2-4 Minggu Sebelum | -87809.598 | 54367.949 | 084 | -1.615 | .111 | -178419.203 | 2800.007 | .366 | 188 | 061 | .521 | 1.92 | | |
| Waktu_Pembelian_Tiket= > 1 Bulan Sebelum | -177877.068 | 77212.567 | 099 | -2.304 | .024 | -306559.506 | -49194.630 | .028 | 264 | 086 | .768 | 1.3 | | |

(Tabel 4.6: Uji Multikolinearitas Coefficients)

Dalam uji multikolinearitas dasar pengambilan keputusan yaitu tidak terjadi gejala multikolinearitas, apabila nilai Tolerance > 0,100 dan nilai VIF < 10,00. Maka, dapat disimpulkan dari tabel 6 tidak ada gejala multikolinearitas karena semua variabel memiliki nilai Tolerance > 0,100 dan nilai VIF < 10,00.

4.2.3.2 Uji Normalitas



(Gambar 4.4: Uji Normalitas)

Pada scatter plot garis regresi pada gambar 4 hasil pengujian asumsi dapat disimpulkan berdistribusi tidak normal karena titik - titik mengikuti garis linear. Namun titik - titik membentuk pola - pola tertentu tidak segaris dengan garis linear, Untuk memastikan apakah kesimpulan dari uji normalitas ini benar. diperlukan pengujian dengan metode kolmogorov-smirnov test.

One-Sample Kolmogorov-Smirnov Test

Unstandardiz ed Residual N 80 Normal Parameters a,b Mean .0000000 Std. Deviation 162082.8628 Most Extreme Differences Absolute .100 Positive .100 Negative -.081 .100 Test Statistic .048° Asymp. Sig. (2-tailed)

(Tabel 4.7: Uji Normalitas One-Sample Kolmogorov-Smirnov Test)
Hasil pengujian Kolmogorov Smirnov didapatkan p-value sebesar 0.048, yaitu lebih kecil dari 0.1. Artinya, gagal tolak Ho atau error berdistribusi normal. Maka, asumsi error normal terpenuhi.

4.2.3.3 Uji Autokorelasi

Model Summarv^b

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin- Watson |
|-------|-------|----------|----------------------|-------------------------------|-------------------|
| 1 | .949ª | .900 | .889 | 170970.614 | 1.995 |

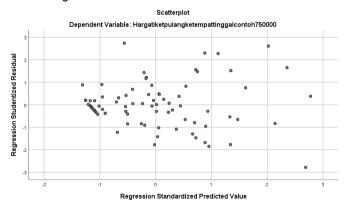
- a. Predictors: (Constant), Waktu_Pembelian_Tiket=> 1 Bulan Sebelum, Transportasi=Pesawat, Transportasi=Bus, Musim=Holiday, Waktu_Pembelian_Tiket=2-4 Minggu Sebelum, Kelas_Transportasi=Eksekutif/Bisnis, Jarak_Tempuh, Waktu_Tempuh
- b. Dependent Variable: Harga Tiket

(Tabel 4.8: Uji Autokorelasi Model Summary)

menurut Field (2009:236) menyatakan "Finally, if you requested the Durbin-Watson statistic it will be found in the last column of the table in SPSS Output 7.5. This statistic informs us about whether the assumption of independent errors is tenable (see section 7.6.2.1). As a conservative rule I suggested that values less than 1 or greater than 3 should definitely raise alarm bells (although I urge you to look up precise values for the situation of interest). The closer to 2 that the value is, the better, and for these data the value is 1.950, which is so close to 2 that the assumption has almost certainly been met."

Berdasarkan uraian dari literatur tersebut, untuk menguji asumsi independensi dari error atau autokorelasi, dapat digunakan uji Durbin-Watson. Nilai statistik dari uji Durbin-Watson yang lebih kecil dari 1 atau lebih besar dari 3 diindikasi terjadi autokorelasi. Hasil uji Durbin-Watson dapat dilihat pada Tabel 7, Tabel Model Summary, pada kolom Durbin-Watson. Diketahui nilai statistik Durbin-Watson adalah 1,995, yakni 1 < 1,995 < 3, maka uji asumsi independensi error autokorelasi terpenuhi.

4.2.3.4 Uji Heteroskedastisitas



(Gambar 4.5 ScatterPlot)

Dalam uji Heteroskedastisitas dasar pengambilan keputusan yaitu tidak terjadi gejala multikolinearitas, apabila titik-titik menyebar di atas dan di bawah angka 0 pada sumbu Y. Maka, dapat disimpulkan dari gambar 4.5 tidak ada gejala Heteroskedastisitas

karena titik-titik menyebar di atas dan di bawah angka 0 pada sumbu Y.

4.2.4 Model Regresi Variable Selection

Untuk mendapatkan model regresi akhir yang baik maka kita menggunakan selection untuk menghilangkan variabel yang memiliki P-Value lebih dari alpha (10%). Kami menggunakan metode *backward selection* untuk mendapatkan model regresi akhir yang baik.

| Coefficients | | | | | | | | | | | | | |
|--------------|---|------------------------------|------------|------|---------------------------------|------|--------------|-------------|------------|-------------------------|------|-----------|-------|
| | | Standardized Coefficients | | | 90.0% Confidence Interval for B | | Correlations | | | Collinearity Statistics | | | |
| Model | | В | Std. Error | Beta | t | Sig. | Lower Bound | Upper Bound | Zero-order | Partial | Part | Tolerance | VIF |
| 5 | (Constant) | -27315.245 | 34323.143 | | 796 | .429 | -84477.878 | 29847.389 | | | | | |
| | Jarak_Tempuh | 534.017 | 44.179 | .579 | 12.088 | .000 | 460.441 | 607.594 | .856 | .813 | .450 | .604 | 1.657 |
| | Transportasi=Pesawat | 585636.594 | 57329.337 | .548 | 10.215 | .000 | 490158.862 | 681114.326 | .765 | .763 | .380 | .481 | 2.078 |
| | Kelas_Transportasi=Eks ekutif/Bisnis | 231644.745 | 49667.496 | .208 | 4.664 | .000 | 148927.237 | 314362.253 | 079 | .474 | .174 | .695 | 1.440 |
| | Waktu_Pembelian_Tiket= > 1 Bulan Sebelum | -129742.214 | 69263.748 | 072 | -1.873 | .065 | -245095.817 | -14388.611 | .028 | 211 | 070 | .939 | 1.065 |

(Tabel 4.9: Model Regresi Variabel Selection Coefficients)

Terdapat 4 variabel yang berpengaruh signifikan, yaitu Jarak Tempuh, Transportasi=Pesawat, Kelas Transportasi=Eksekutif/Bisnis, dan Waktu Pembelian Tiket=>1 Bulan Sebelum. Masing masing memiliki nilai p-value kurang dari taraf signifikansi 0.1. Model regresi dapat dinyatakan sebagai:

 $Y = -27315.245 + 534.017(Jarak\ Tempuh) + \\ 585636.594(Transportasi:\ Pesawat) + \\ 231644.745(Kelas\ Transportasi:\ Eksekutif\ /Bisnis) - \\ 129742.214(Waktu\ Pembelian\ Tiket:\ > 1\ Bulan\ Sebelum)$

4.2.5 Kebaikan Model Regresi Akhir Model Summary^f

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin- Watson |
|-------|-------------------|----------|----------------------|-------------------------------|-------------------|
| 1 | .949 ^a | .900 | .889 | 170970.614 | |
| 2 | .949 ^b | .900 | .890 | 169792.306 | |
| 3 | .949° | .900 | .892 | 168740.949 | |
| 4 | .948 ^d | .899 | .892 | 168528.949 | |
| 5 | .947 ^e | .896 | .891 | 169658.599 | 1.998 |

(Tabel 4.10: Kebaikan Model Regresi Akhir Model Summary)

Pada tabel 10, model yang kita gunakan adalah nomor 5. Nilai R-Squared yang didapatkan adalah 0.896. Artinya, variabel prediktor mampu menjelaskan variansi variabel respon sebesar 89.6%, sedangkan sisanya yaitu 10.4% dijelaskan oleh variabel lain yang tidak diteliti. Secara keseluruhan, model ini memiliki tingkat kebaikan yang tinggi karena mampu menjelaskan sebagian besar variabilitas variabel respon yang diamati.

BAB V KESIMPULAN

5.1 Kesimpulan

Berdasarkan hasil survey yang didapat, data yang terkumpul sebanyak 80 responden mahasiswa prodi Teknologi Sains Data, Universitas Airlangga, dengan analisis regresi yang sudah kami lakukan untuk meneliti bagaimana pola hubungan antara variabel prediktor(X) terhadap variabel respon berupa Harga Tiket (Y), didapatkan hasil sebagai berikut:

1. Analisis eksplorasi data

Dari hasil analisis eksplorasi data yang dilakukan didapatkan dari grafik histogram bersifat right-skewness yang menunjukkan adanya data yang lebih ekstrim atau langka di sisi kanan distribusi yaitu ada pada data rentang 1.500.000 - 2.000.000 Rupiah, Hal tersebut menjadikan visualisasi data pada boxplot memiliki outlier yang berada di harga 2.000.000 sehingga diduga bahwa variabel respon tidak berdistribusi normal.

2. Analisis Regresi Linier Berganda

a. Regresi Variabel Dependen dan Independen
 Dari hasil analisis didapatkan nilai R-Squared sebesar 0,90.
 Sehingga model ini memiliki tingkat kebaikan yang tinggi karena mampu menjelaskan variansi variabel respon sebesar 90%.

b. Uji signifikansi Parameter

a. Uji Serentak

Didapatkan hasil nilai Fhitung 80.039 dengan P-Value 0.0000 atau kurang dari taraf signifikansi 0,1 artinya variabel independen secara bersama-sama serentak mempunyai pengaruh yang signifikan terhadap variabel dependen (Hargatiketpulangtempattinggal)

b. Uji Parsial

Didapatkan hasil bahwa sebanyak 4 variabel yang signifikan yaitu Jarak_Tempuh, Transportasi=Pesawat, Kelas_Transportasi=Eksekutif/Bisnis,Waktu Pembelian Tiket =>1 Bulan Sebelum memiliki P-Value kurang dari taraf signifikansi 0,1. Sedangkan sisanya tidak terdapat hubungan signifikan dengan variabel dependen (Hargatiketpulangtempattinggal)

c. Uji Asumsi Error

- Pada uji Multikolinearitas dapat disimpulkan tidak ada gejala multikolinearitas karena semua variabel memiliki nilai Tolerance > 0,100 dan nilai VIF < 10,00.
- Pada uji Normalitas hasil pengujian asumsi dapat disimpulkan berdistribusi kurang normal karena titik - titik

- membentuk pola tertentu, namun masih mengikuti garis linear
- Pada uji Autokorelasi diketahui nilai statistik Durbin-Watson adalah 1,995, yakni di antara 1 dan 3, 1 < 1,995 < 3, maka uji asumsi independensi error autokorelasi terpenuhi.
- Pada uji Heteroskedastisitas diketahui titik-titik menyebar di atas dan di bawah angka 0 pada sumbu Y sehingga tidak ada gejala Heteroskedastisitas.
- d. Y = -27315.245 + 534.017(Jarak Tempuh) + 585636.594(Transportasi: Pesawat) + 231644.745(Kelas Transportasi: Eksekutif /Bisnis) 129742.214(Waktu Pembelian Tiket: > 1 Bulan Sebelum) -27315,245 merupakan persentase harga tiket pulang, 534,017% adalah persentase peningkatan harga tiket jika jarak tempuh bertambah, 585.636,594% adalah persentase peningkatan harga tiket jika menggunakan transportasi pesawat, 231.644,745% adalah persentase peningkatan harga tiket jika menggunakan kelas transportasi eksekutif/bisnis, dan -129.742,214% adalah persentase penurunan harga tiket jika waktu pembelian tiket lebih dari 1 bulan sebelumnya.
- e. Dengan $R^2 = 0.9$, dan R-adj 0,889 artinya variabel prediktor mampu menjelaskan variansi variabel respon sebesar 89.6%, Secara keseluruhan, model ini memiliki tingkat kebaikan yang tinggi karena mampu menjelaskan sebagian besar variabilitas variabel respon yang diamati.

5.2 Saran

5.2.1 Saran untuk Pembaca

Dalam membeli tiket pulang ke tempat tinggal hendaknya mempertimbangkan apa saja variabel yang signifikan atau dapat mempengaruhi kenaikan harga tiket. Adapun variabel yang dapat mempengaruhi yaitu, jarak tempuh, transportasi, dan waktu pembelian tiket.

5.2.2 Saran untuk Penelitian

Adapun beberapa saran yang dapat bermanfaat sebagai acuan dan membangun pada penelitian kami atau pada peneliti yang ingin mengembangkan selanjutnya adalah sebagai berikut :

- 1. Mencari data responden lebih banyak agar dapat menggunakan alpha 0.05 pada analisis yang akan dilakukan.
- 2. Menentukan variabel prediktor dengan teliti yang berpeluang memiliki korelasi dengan variabel dependen agar R-Square dan Adjusted R-Square bisa meningkat lebih dari 50%.

DAFTAR PUSTAKA

Larasati, R. A. (2020). Pola konsumsi mahasiswa pulang kampung dan masyarakat pada pandemi covid-19 di Kota Bandung. *Jambura Economic Education Journal*, 2(2), 90-99.

Setiana B. (2015). PRINSIP-PRINSIP POKOK PENGELOLAAN JASA TRANSPORTASI UDARA. *Jurnal Ilmiah WIDYA*,103-109.(definisi Transportasi ama pesawat)

Humam Santosa Utomo. Manajemen Transportasi, Malang: Pascasarjana Universitas Brawijaya, 2010.

Haryono Sukarto. Transportasi Perkotaan dan Lingkungan, Jurnal Teknik, Jakarta. 2006.

Khoiriyatun, Emi. Yona Sidratul Munti, Novi. Gusman, Deddy. (2021) Estimasi Jarak Dan Waktu Bus Sekolah Pada System Pemetaan Trayek Bus Sekolah Dan Halte Di Central Business District (CBD) Bangkinang, 1-2. (jarak dan waktu)

Sriastuti, Dewa Ayu Nyoman. (2015). Kereta Api Pilihan Utama Sebagai Moda Alternatif Angkutan Umum Massal, 26-34. (kereta api)

Prastyo, Adhitia. Fadlan, Fadlan; Fadjriani, Lia. (2020). Analisis Yuridis Terhadap Keberangkatan Kapal Penumpang Tanpa Adanya Surat Persetujuan Berlayar (Studi Penelitian Kantor Kesyahbandaran Dan Otoritas Pelabuhan Khusus Batam), 1-15. (kapal)

Silalahi, Novrika. 2020. Analisis Regresi Berganda Dengan Variabel Dummy Pada Faktor Kejadian Infeksi Saluran Pernafasan Akut Pada Remaja Terhadap Kondisi Fisik Rumah

Field, A. 2009. Discovering Statistics Using SPSS, 3rd Edition. London: Sage.

LAMPIRAN

Sintaks Mencari Histogram dan Boxplot

```
# Histogram dan Boxplot
sns.set(rc={'figure.figsize':(6,4)})
sns.distplot(df['Harga Tiket'], bins=10,
kde_kws={'linewidth': 3})
plt.show()

plt.figure(figsize=(6,4))
sns.boxplot(x=df['Harga Tiket'], color='grey')
plt.xlabel('Harga Tiket')
plt.title('Boxplot Harga Tiket Pulang')
plt.show
```

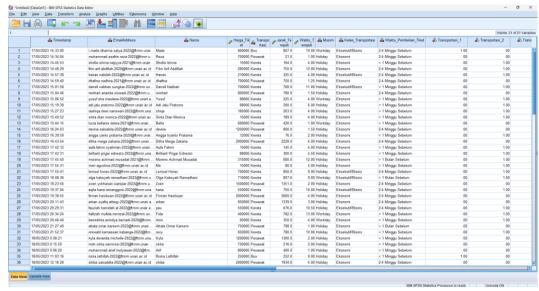
Sintaks Mencari Frekuensi Penyebaran Data Seluruh Variabel

```
plt.figure(figsize=(18, 7))

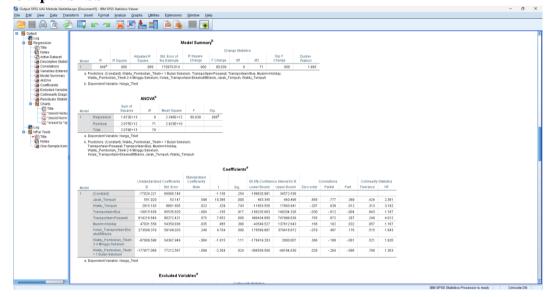
predictors = df.drop(columns=['Harga Tiket']).columns
response = df['Harga Tiket']

for i, col in enumerate(predictors):
    plt.subplot(1, len(predictors), i+1)
    x = df[col]
    y = response
    plt.scatter(x, y, marker='o', color='darkblue',
alpha=0.5)
    plt.title(col)
    plt.xlabel(col)
    plt.ylabel('Harga Tiket')
    plt.tight_layout()
```

SPSS



Output SPSS



File SPSS (Data dan Output)

Data dan Output SPPS

Google Forms

https://forms.gle/5Cbz33xbyXvMq3JFA



Data set

Mau Tiket Pulang Gratis? (Jawaban)