

PROJECT AKHIR METODE STATISTIKA

**“Pengaruh Aktivitas Digital Mahasiswa Teknologi Sains Data Angkatan 2022
terhadap Biaya Pengeluaran dalam Pembelian Kuota Internet Bulanan”**



Dosen Pengampu :

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

Disusun oleh :

Kelompok C / SD-A2

- | | | |
|----|--------------------------|-----------|
| 1. | Arkan Syafiq Attaqy | 164221062 |
| 2. | Haryo Bismo Wicaksono | 164221095 |
| 3. | Richo Wignyo Aji Saputra | 164221013 |
| 4. | Salma Ayu Hanifah | 164221012 |
| 5. | Verdyan Farrel Biliarsa | 164221076 |

**PROGRAM STUDI TEKNOLOGI SAINS DATA
FAKULTAS TEKNOLOGI MAJU DAN MULTIDISIPLIN
UNIVERSITAS AIRLANGGA**

DAFTAR ISI

| | |
|-----------------------------------------------------------------|-----------|
| DAFTAR ISI..... | i |
| DAFTAR TABEL..... | ii |
| DAFTAR GAMBAR..... | iii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Rumusan Masalah..... | 1 |
| 1.3 Tujuan Penelitian..... | 2 |
| 1.4 Batasan Penelitian..... | 2 |
| BAB II TINJAUAN PUSTAKA..... | 3 |
| 2.1 Tinjauan Statistika..... | 3 |
| 2.1.1 Regresi Linier..... | 3 |
| 2.1.2 Uji Asumsi..... | 3 |
| 2.2 Tinjauan Non-statistika..... | 3 |
| 2.2.1 Data Seluler..... | 3 |
| 2.2.2 Wi-Fi..... | 4 |
| 2.2.3 Provider..... | 4 |
| 2.2.4 Media Sosial..... | 4 |
| 2.2.5 Media Streaming..... | 4 |
| BAB III METODOLOGI PENELITIAN..... | 5 |
| 3.1 Metode Pengambilan Sampel..... | 5 |
| 3.1.1 Populasi..... | 5 |
| 3.1.2 Sampel..... | 5 |
| 3.1.3 Tingkat Signifikansi..... | 6 |
| 3.2 Variabel Penelitian..... | 6 |
| 3.2.1 Variabel Dependen..... | 6 |
| 3.2.2 Variabel Independen..... | 6 |
| BAB IV HASIL DAN ANALISIS PENELITIAN..... | 7 |
| 4.1 Analisis Eksplorasi Data (EDA)..... | 7 |
| 4.1.1 Visualisasi Pola atau Korelasi antar Variabel..... | 7 |
| 4.1.2 Penghitungan Korelasi antar Variabel..... | 7 |
| 4.2 Analisis Regresi Linier Berganda..... | 8 |
| 4.2.1 Regresi Variabel Dependen dengan Variabel Independen..... | 8 |
| 4.2.2 Uji Signifikansi Parameter..... | 9 |
| 4.2.3 Uji Asumsi..... | 10 |
| BAB V PENUTUP..... | 14 |
| 5.1 Kesimpulan..... | 14 |
| 5.2 Saran..... | 15 |
| DAFTAR PUSTAKA..... | 16 |
| LAMPIRAN..... | 17 |

DAFTAR TABEL

| | |
|---------------------|----|
| Tabel 4.1.2 | 7 |
| Tabel 4.2.1 a | 8 |
| Tabel 4.2.1 b | 8 |
| Tabel 4.2.1 c | 9 |
| Tabel 4.2.2 a | 9 |
| Tabel 4.2.2 b | 9 |
| Tabel 4.2.3 a | 10 |
| Tabel 4.2.3 b | 11 |
| Tabel 4.2.3 c | 11 |
| Tabel 4.2.3 d | 12 |
| Tabel 4.2.4 a | 12 |
| Tabel 4.2.4 b | 12 |
| Tabel 4.2.5 a..... | 13 |
| Tabel 4.2.5 b..... | 13 |

DAFTAR GAMBAR

| | |
|----------------------|----|
| Gambar 4.1.1 a | 7 |
| Gambar 4.1.1 b | 7 |
| Gambar 4.1.2 c | 7 |
| Gambar 4.2.3 a | 10 |
| Gambar 4.2.3 b | 10 |
| Gambar 4.2.3 c | 11 |
| Gambar 4.2.3 d | 12 |

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era digital saat ini, teknologi informasi dan komunikasi memainkan peran yang sangat penting dalam kehidupan sehari-hari, termasuk di kalangan mahasiswa. Peningkatan penggunaan perangkat mobile, akses internet yang lebih mudah, dan perkembangan aplikasi online telah membawa perubahan besar dalam cara mahasiswa mengakses informasi, belajar, dan berkomunikasi.

Biaya pengeluaran dalam pembelian kuota internet bulanan menjadi isu yang signifikan bagi mahasiswa Teknologi Sains Data Angkatan 2022. Ketergantungan mereka pada akses internet yang luas dan kuota data yang cukup untuk mengakses sumber daya digital yang relevan dengan studi mereka menjadi faktor penentu yang mempengaruhi pengeluaran bulanan mereka. Namun, meskipun pentingnya kuota internet dalam kehidupan mahasiswa Teknologi Sains Data, belum ada studi yang secara khusus menginvestigasi pengaruh aktivitas digital mereka terhadap biaya pengeluaran mereka. Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi dan menganalisis pengaruh aktivitas digital mahasiswa Teknologi Sains Data Angkatan 2022 terhadap biaya pengeluaran mereka dalam pembelian kuota internet bulanan.

Melalui penelitian ini, diharapkan dapat diperoleh pemahaman yang lebih baik tentang pola penggunaan internet oleh mahasiswa Teknologi Sains Data dan dampaknya terhadap pengeluaran mereka. Hasil penelitian ini diharapkan dapat memberikan wawasan kepada mahasiswa, lembaga pendidikan, dan penyedia layanan internet tentang pentingnya pengelolaan biaya penggunaan internet dan pengembangan paket kuota yang sesuai dengan kebutuhan mahasiswa. Dengan demikian, penelitian ini memiliki kepentingan yang signifikan dalam konteks pengaruh aktivitas digital mahasiswa Teknologi Sains Data angkatan 2022 terhadap biaya pengeluaran dalam pembelian kuota internet bulanan.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah di atas, maka rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana pengaruh aktivitas digital mahasiswa Teknologi Sains Data angkatan 2022 terhadap biaya pengeluaran dalam pembelian kuota internet bulanan?

2. Bagaimana model regresi linier berganda yang baik terhadap biaya pengeluaran dalam pembelian kuota internet bulanan mahasiswa Teknologi Sains Data angkatan 2022?

1.3 Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah di atas, maka tujuan penelitian ini adalah sebagai berikut:

1. Untuk mengetahui pengaruh aktivitas digital mahasiswa Teknologi Sains Data angkatan 2022 terhadap biaya pengeluaran dalam pembelian kuota internet bulanan
2. Untuk mendapatkan model regresi linier berganda yang baik terhadap biaya pengeluaran dalam pembelian kuota internet bulanan mahasiswa Teknologi Sains Data angkatan 2022

1.4 Batasan Penelitian

Batasan penelitian yang dilakukan yaitu melakukan survei untuk mahasiswa Teknologi Sains Data Universitas Airlangga dengan kriteria berikut:

1. Mahasiswa angkatan tahun 2022 atau semester 2 di tahun 2023.
2. Memiliki minimal responden 89 mahasiswa.
3. Berlangganan kuota internet tiap bulannya.

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Statistika

2.1.1 Regresi Linier

Regresi linear adalah suatu pendekatan statistik yang digunakan untuk memodelkan hubungan linier antara variabel dependen (variabel yang ingin diprediksi) dan satu atau lebih variabel independen (variabel yang digunakan untuk memprediksi). Pada regresi linear, kita mencari persamaan garis lurus terbaik yang menggambarkan hubungan antara variabel dependen dan independen. Persamaan tersebut dapat digunakan untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen yang diketahui.

2.1.2 Uji Asumsi

Uji asumsi adalah proses statistik yang digunakan untuk memeriksa apakah data yang digunakan dalam analisis statistik memenuhi asumsi yang diperlukan untuk metode analisis tertentu. Asumsi ini adalah syarat-syarat atau hipotesis yang harus terpenuhi agar hasil analisis statistik dapat dianggap valid. Uji asumsi dilakukan sebelum atau setelah melakukan analisis statistik untuk memastikan bahwa data yang digunakan memenuhi asumsi yang diperlukan. Jika asumsi tidak terpenuhi, maka mungkin perlu dilakukan transformasi data atau menggunakan metode analisis yang lebih sesuai. Ada beberapa contoh asumsi yang umum dalam analisis statistik meliputi:

1. Asumsi normalitas
2. Asumsi variansi error konstan (homoskedastisitas)
3. Asumsi independensi error (autokorelasi)
4. Asumsi linearitas
5. Multikolinieritas

2.2 Tinjauan Non-statistika

2.2.1 Data Seluler

Data seluler adalah layanan yang memungkinkan pengguna untuk mengakses internet dan mentransfer data melalui jaringan seluler. Dalam jurnal online "The Evolution of Cellular Data Technologies" oleh Johnson et al. (2019), peneliti menjelaskan perkembangan dan inovasi dalam teknologi data seluler, termasuk transisi dari 3G ke 4G dan kemudian ke 5G.

2.2.2 Wi-Fi

WiFi adalah teknologi nirkabel yang memungkinkan pengiriman data melalui gelombang radio. Ini telah mengubah cara kita terhubung dengan internet dan mengakses informasi dengan mudah. Dalam jurnal online "The Impact of WiFi Technology on Society" oleh Smith et al. (2020), mereka menjelaskan bagaimana WiFi telah mempengaruhi berbagai aspek kehidupan kita, termasuk komunikasi, pendidikan, bisnis, dan hiburan. Penelitian ini menunjukkan bahwa kehadiran WiFi telah meningkatkan konektivitas global dan memberikan akses lebih mudah ke sumber daya informasi. WiFi juga telah memungkinkan inovasi baru seperti Internet of Things (IoT).

2.2.3 Provider

Provider internet atau penyedia layanan internet adalah perusahaan atau organisasi yang menyediakan akses internet kepada pengguna melalui infrastruktur dan jaringan yang mereka miliki. Peran utama dari provider internet adalah menyediakan koneksi internet kepada pengguna, provider internet biasanya menawarkan berbagai paket layanan yang disesuaikan dengan kebutuhan pengguna. Paket ini mencakup kecepatan akses internet, kuota data, dan biaya berlangganan yang berbeda-beda. Pada penelitian ini kami mengerucutkan penelitian kami terhadap 4 provider besar di Indonesia, keempat provider tersebut adalah Telkomsel, XL Axiata, Indosat, dan Smartfren.

2.2.4 Media Sosial

Boyd dalam Nasrullah (2015) media sosial sebagai kumpulan perangkat lunak yang memungkinkan individu maupun komunitas untuk berkumpul, berbagi, berkomunikasi, dan dalam kasus tertentu saling berkolaborasi atau bermain. Media sosial merujuk kepada platform-platform online yang memungkinkan pengguna untuk berinteraksi, berbagi konten, dan berkomunikasi dengan pengguna lain secara virtual.

2.2.5 Media Streaming

Media streaming telah mengubah cara kita mengonsumsi konten multimedia. Dalam jurnal online "The Impact of Media Streaming on Entertainment Consumption" (2020), penelitian ini membahas peran media streaming dalam mengubah pola konsumsi hiburan. Penelitian ini menunjukkan bahwa media streaming memberikan kemudahan akses, variasi konten, dan fleksibilitas waktu dalam menonton film, acara TV, dan musik. Hal ini telah mengubah industri hiburan dan memberikan pengalaman pengguna yang lebih personal.

BAB III

METODOLOGI PENELITIAN

3.1 Metode Pengambilan Sampel

3.1.1 Populasi

Populasi adalah gabungan dari seluruh elemen yang terbentuk peristiwa, hal, atau orang yang memiliki karakteristik serupa yang menjadi pusat perhatian peneliti, karena dipandang sebagai semesta penelitian. Dalam penelitian ini, populasi yang digunakan adalah mahasiswa program studi Teknologi Sains Data (TSD) angkatan 2022, Universitas Airlangga dengan jumlah estimasi mahasiswa adalah sebanyak 113 mahasiswa.

3.1.2 Sampel

Sampel adalah bagian dari jumlah dan karakteristik yang dimiliki oleh populasi. Ada dua teknik *sampling* (pengambilan sampel) yaitu *probability sampling* dan *non probability sampling*. *Probability sampling* adalah teknik pengambilan sampel yang memberikan setiap elemen (anggota) populasi kesempatan yang sama untuk dipilih sebagai anggota sampel. Sedangkan *non-probability sampling* adalah teknik pengambilan sampel yang tidak memberikan peluang/peluang yang sama untuk setiap elemen anggota populasi untuk dijadikan sampel.

Metode pengambilan sampel yang digunakan dalam penelitian ini adalah *non probability sampling* yaitu *convenience/accidental sampling*, di mana sampel terpilih karena berada pada waktu dan tempat yang tepat. Sampel diambil dengan cara menyebar kuesioner via google form di grup WhatsApp. Untuk menentukan jumlah sampel yang akan digunakan dalam penelitian, peneliti menggunakan rumus Slovin.

Adapun rumus yang digunakan :

$$n = \frac{N}{1 + Ne^2}$$

Keterangan :

n : ukuran sampel

N : ukuran populasi

e : nilai margin of error

Perhitungan menggunakan rumus Slovin :

$$n = \frac{N}{1 + Ne^2}$$

$$n = \frac{113}{1 + 113(0,05)^2}$$

$$n = 88,11 \approx 89$$

Sehingga dapat disimpulkan bahwa dalam penelitian ini dibutuhkan paling sedikit 89 responden. Kemudian dari hasil penyebaran kuesioner didapatkan responden sebanyak 89 mahasiswa.

3.1.3 Tingkat Signifikansi

Dalam penelitian ini, tingkat signifikansi yang digunakan adalah 5%. Tingkat tersebut dipilih rendah agar menunjukkan tingkat kepercayaan yang lebih tinggi dalam hasil yang dihasilkan. Dengan menggunakan tingkat signifikansi ini, diharapkan penelitian dapat dilakukan secara efisien tanpa mengorbankan keakuratan hasil.

3.2 Variabel Penelitian

3.2.1 Variabel Dependen

Variabel dependen adalah variabel yang dipengaruhi oleh variabel independen. Dalam penelitian ini, variabel dependennya adalah biaya pengeluaran dalam pembelian kuota internet bulanan.

3.2.2 Variabel Independen

Variabel independen adalah variabel yang berperan sebagai penyebab timbulnya variabel dependen. Dalam penelitian ini, variabel independennya terdiri dari:

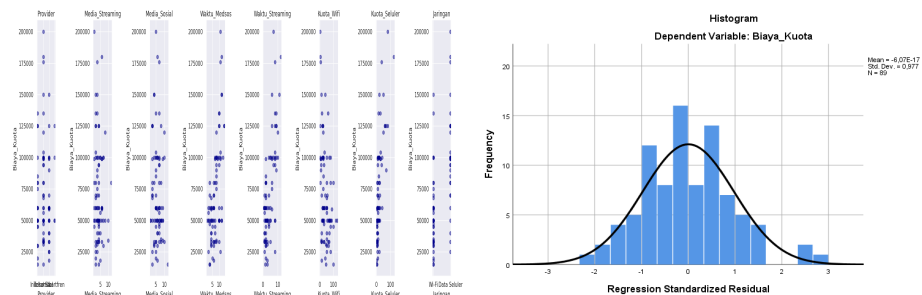
- Jenis provider
- Jumlah media sosial
- Jumlah aplikasi/web streaming
- Waktu penggunaan media sosial per hari (dalam jam)
- Waktu penggunaan aplikasi/web streaming per minggu (dalam jam)
- Jumlah penggunaan data wi-fi
- Jumlah penggunaan data seluler
- Jaringan internet yang sering digunakan

BAB IV

HASIL DAN ANALISIS PENELITIAN

4.1 Analisis Eksplorasi Data (EDA)

4.1.1 Visualisasi Pola atau Korelasi antar Variabel



Gambar 4.1.1 a) Scatter Plot Korelasi

Gambar 4.1.1. b) Histogram Biaya Kuota Bulanan

Di atas merupakan visualisasi histogram dan scatterplot korelasi antara variabel dependen dengan variabel independen. Histogram pada variabel respon biaya pengeluaran dalam pembelian kuota internet bulanan berbentuk bell shaped, sehingga dapat disimpulkan data berdistribusi normal.

4.1.2 Penghitungan Korelasi antar Variabel



Gambar 4.1.2 c) Visualisasi Heatmap

| Correlations | | | | | | | | | | | | |
|---------------------|--------------------|-------------|-----------------|--------------|--------------|-----------------|------------|---------------|--------------------|--------------------|-------------|----------------|
| | | Biaya_Kuota | Media_Streaming | Media_Sosial | Waktu_Medsos | Waktu_Streaming | Kuota_Wifi | Kuota_Seluler | Provider=Smartfren | Provider=Telkomsel | Provider=XL | Jaringan=Wi-Fi |
| Pearson Correlation | Biaya_Kuota | 1,000 | -.071 | -.101 | .594 | .701 | -.390 | .736 | .060 | .150 | -.111 | -.591 |
| | Media_Streaming | -.071 | 1,000 | .492 | .080 | -.074 | .050 | -.049 | -.064 | .224 | -.087 | .156 |
| | Media_Sosial | -.101 | .492 | 1,000 | .095 | -.116 | .002 | -.121 | -.146 | .276 | .023 | .138 |
| | Waktu_Medsos | .594 | .080 | .095 | 1,000 | .649 | -.282 | .510 | .029 | .009 | .012 | -.499 |
| | Waktu_Streaming | .701 | -.074 | -.116 | .649 | 1,000 | -.143 | .638 | -.018 | .084 | .017 | -.532 |
| | Kuota_Wifi | -.390 | .050 | .002 | -.282 | -.143 | 1,000 | -.324 | -.051 | .077 | -.160 | .439 |
| | Kuota_Seluler | .736 | -.049 | -.121 | .510 | .638 | -.324 | 1,000 | .064 | -.005 | -.030 | -.553 |
| | Provider=Smartfren | .060 | -.064 | -.146 | .029 | -.018 | -.051 | .064 | 1,000 | -.196 | -.128 | .050 |
| | Provider=Telkomsel | .150 | .224 | .276 | .009 | .084 | .077 | -.005 | -.196 | 1,000 | -.533 | -.029 |
| | Provider=XL | -.111 | -.087 | .023 | .012 | .017 | -.160 | -.030 | -.128 | -.533 | 1,000 | -.039 |
| | Jaringan=Wi-Fi | -.591 | .156 | .138 | -.499 | -.532 | .439 | -.553 | .050 | -.029 | -.039 | 1,000 |

Tabel 4.1.2 Perhitungan Nilai Korelasi SPSS

Di atas merupakan nilai korelasi sebelum dilakukannya analisis regresi. Visualisasi heatmap dan perhitungan menggunakan spss di atas menampilkan nilai korelasi antar variabel yang didapatkan dengan menggunakan metode korelasi pearson. Berdasarkan nilai korelasi, dapat diketahui bahwa variabel prediktor yang berkorelasi positif atau berpengaruh terhadap variabel respons (biaya pengeluaran dalam

pembelian kuota internet bulanan) adalah kuota seluler (0.74), waktu streaming (0.7), dan waktu medsos (0.59).

4.2 Analisis Regresi Linier Berganda

4.2.1 Regresi Variabel Dependen dengan Variabel Independen

- a) Regresi variabel Independen dengan variabel Dependen **sebelum** dilakukan backwards elimination:

| OLS Regression Results | | | | | | |
|------------------------|------------------|---------------------|----------|-------|-----------|----------|
| Dep. Variable: | Biaya_Kuota | R-squared: | 0.707 | | | |
| Model: | OLS | Adj. R-squared: | 0.670 | | | |
| Method: | Least Squares | F-statistic: | 18.84 | | | |
| Date: | Sat, 20 May 2023 | Prob (F-statistic): | 5.07e-17 | | | |
| Time: | 13:49:41 | Log-Likelihood: | -1011.5 | | | |
| No. Observations: | 89 | AIC: | 2045. | | | |
| Df Residuals: | 78 | BIC: | 2072. | | | |
| Df Model: | 10 | | | | | |
| Covariance Type: | nonrobust | | | | | |
| | coef | std err | t | P> t | [0.025 | 0.975] |
| const | 4.616e+04 | 1.16e+04 | 3.984 | 0.000 | 2.31e+04 | 6.92e+04 |
| Media_Streaming | -644.9278 | 1317.437 | -0.490 | 0.626 | -3267.743 | 1977.887 |
| Media_Sosial | -581.2537 | 1442.508 | -0.403 | 0.688 | -3453.067 | 2290.560 |
| Waktu_Medsos | 1491.2423 | 1065.440 | 1.400 | 0.166 | -629.886 | 3612.370 |
| Waktu_Streaming | 4016.8573 | 1310.092 | 3.066 | 0.003 | 1408.664 | 6625.051 |
| Kuota_Wifi | -231.5289 | 96.643 | -2.396 | 0.019 | -423.931 | -39.127 |
| Kuota_Seluler | 683.4313 | 151.809 | 4.524 | 0.000 | 361.295 | 1005.567 |
| Provider_Smartfren | 7910.6060 | 1.23e+04 | 0.646 | 0.520 | -1.65e+04 | 3.23e+04 |
| Provider_Telkomsel | 9704.2170 | 6451.197 | 1.504 | 0.137 | -3139.030 | 2.25e+04 |
| Provider_Xl | -5901.7673 | 6996.978 | -0.843 | 0.402 | -1.98e+04 | 8028.146 |
| Jaringan_Wi-Fi | -6327.4388 | 6881.085 | -0.920 | 0.361 | -2e+04 | 7371.749 |

Tabel 4.2.1 a) OLS Regresi Awal

Persamaan Regresi yang didapat:

$$\text{biaya_kuota} = 4.616e+04 - 644.9278 (\text{media_streaming}) - 581.2537 (\text{media_sosial}) + 1491.2423 (\text{waktu_medsos}) + 4016.8573 (\text{waktu_streaming}) - 231.5289 (\text{kuota_wifi}) + 683.4313 (\text{kuota_seluler}) + 7910.6060 (\text{provider_smartfren}) + 9704.2170 (\text{provider_telkomsel}) - 5901.7673 (\text{provider_xl}) - 6327.4388 (\text{jaringan_wi-fi})$$

R-Square yang didapat dari model regresi awal:

| | Model | R-Square | Adjusted R-Square |
|---|-------------------------------|----------|-------------------|
| 0 | Regression + Dummy Encoding | 0.707178 | 0.669637 |
| 1 | Regression + One-Hot Encoding | 0.707178 | 0.669637 |

Tabel 4.2.1 b) R-Square Regresi Awal

Model Regresi Awal memiliki R-Square sebesar 0.707178, yang menunjukkan bahwa model mampu menjelaskan sekitar 70.71% dari variasi dalam variabel target (biaya_kuota). Model Regresi Awal memiliki Adjusted R-Square sebesar 0.669637, yang menunjukkan bahwa model mampu menjelaskan sekitar 66.96% dari variasi dalam variabel target setelah mempertimbangkan jumlah variabel independen yang digunakan dalam model.

Secara keseluruhan, kedua model menunjukkan kinerja yang baik dalam menjelaskan hubungan antara variabel independen dengan variabel dependen, dan penggunaan encoding tertentu tidak mempengaruhi kinerja model dalam hal ini.

- b) Regresi variabel Independen dengan variabel Dependen **setelah** dilakukan backwards elimination:

| Coefficients ^a | | | | | | | | |
|---------------------------|-----------------|-----------------------------|------------|---------------------------|--------|------|-------------------------|-------|
| | | Unstandardized Coefficients | | Standardized Coefficients | | | Collinearity Statistics | |
| Model | | B | Std. Error | Beta | t | Sig. | Tolerance | VIF |
| 1 | (Constant) | 48616.569 | 6173.243 | | 7.875 | .000 | | |
| | Waktu_Streaming | 5614.814 | 1055.181 | .424 | 5.321 | .000 | .587 | 1.705 |
| | Kuota_Seluler | 717.426 | 154.986 | .388 | 4.629 | .000 | .532 | 1.880 |
| | Kuota_Wifi | -302.145 | 88.061 | -.226 | -3.431 | .001 | .858 | 1.166 |
| | Provider=XL | -12534.158 | 5484.734 | -.142 | -2.285 | .025 | .963 | 1.038 |

a. Dependent Variable: Biaya_Kuota

Tabel 4.2.1 c) OLS Regresi setelah backwards elimination

Persamaan regresi dari output tersebut adalah:

$$\text{biaya_kuota} = 48616.569 + 5614.814 (\text{waktu_streaming}) + 717.426 (\text{kuota_seluler}) - 302.145 (\text{kuota_wifi}) - 12534.158 (\text{provider_xl})$$

4.2.2 Uji Signifikansi Parameter

a. Uji Serentak

| | F-hitung | p-value | alpha | Kesimpulan |
|---|-----------|--------------|-------|------------|
| 0 | 46.037083 | 2.008259e-20 | 0.05 | Tolak H0 |

Tabel 4.2.2. a) Uji Serentak

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1 : \text{Minimal ada satu } \beta_n \neq 0$$

Kesimpulan:

Berdasarkan perhitungan diatas, dapat kita lihat bahwa nilai p-value (2.008259e-20) lebih kecil dari nilai alpha (0.05) yang artinya kita dapat simpulkan Tolak H0. Berarti terdapat hubungan antara variabel independen dengan variabel dependen.

b. Uji Parsial

| | Variabel | t-hitung | p-value | alpha | Kesimpulan |
|---|-----------------|-----------|--------------|-------|------------|
| 0 | const | 7.875369 | 1.076670e-11 | 0.05 | Tolak H0 |
| 1 | Waktu_Streaming | 5.321186 | 8.417195e-07 | 0.05 | Tolak H0 |
| 2 | Kuota_Wifi | -3.431095 | 9.347256e-04 | 0.05 | Tolak H0 |
| 3 | Kuota_Seluler | 4.628957 | 1.325087e-05 | 0.05 | Tolak H0 |
| 4 | Provider_XL | -2.285281 | 2.481609e-02 | 0.05 | Tolak H0 |

Tabel 4.2.2. b) Uji Parsial

Hipotesis:

$$H_0: \beta_1 = 0 \quad H_0: \beta_2 = 0 \quad H_0: \beta_3 = 0 \quad H_0: \beta_4 = 0$$

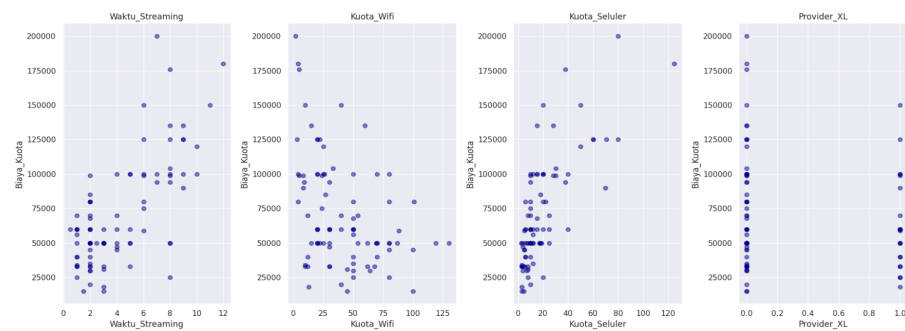
$$H_1: \beta_1 \neq 0 \quad H_1: \beta_2 \neq 0 \quad H_1: \beta_3 \neq 0 \quad H_1: \beta_4 \neq 0$$

Kesimpulan:

- Variabel waktu_streaming, kuota_wifi, kuota_seluler, provider_xl mengalami Tolak H0, yang artinya variabel-variabel tersebut berpengaruh terhadap biaya pengeluaran dalam pembelian kuota internet bulanan.

4.2.3 Uji Asumsi

a) Uji Linearitas



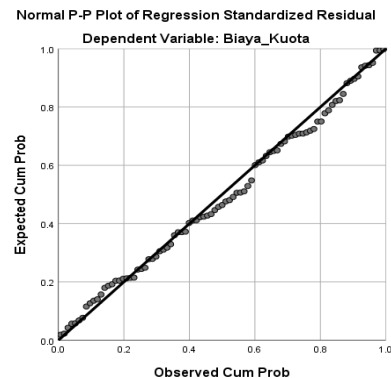
Gambar 4.2.3. a) Scatterplot Uji Linearitas

Dapat diamati pada scatter plot diatas (Hubungan setiap variabel independen x setelah dilakukan *backward elimination* dengan variabel y) menunjukkan bahwa variabel waktu_streaming memiliki hubungan yang linier dengan biaya pengeluaran dalam pembelian kuota internet bulanan. Sedangkan pada variabel kuota_wi-fi, kuota_seluler, dan provider_xl tidak memiliki hubungan yang linier dengan biaya pengeluaran dalam pembelian kuota internet bulanan.

b) Uji Normalitas

| One-Sample Kolmogorov-Smirnov Test | | |
|------------------------------------|---------------------|-------------------------|
| | | Unstandardized Residual |
| N | 89 | |
| Normal Parameters ^{a,b} | Mean | .0000000 |
| | Std. Deviation | 22382.70765 |
| Most Extreme Differences | Absolute | .075 |
| | Positive | .075 |
| | Negative | -.070 |
| Test Statistic | .075 | |
| Asymp. Sig. (2-tailed) | .200 ^{c,d} | |

a. Test distribution is Normal.
b. Calculated from data.
c. Lilliefors Significance Correction.
d. This is a lower bound of the true significance.



Tabel 4.2.3. a) Kolmogorov - Smirnov Test

Gambar 4.2.3. b) Plot Distribusi Normal

- Pengujian dengan grafik

Sebaran data berada di sekitar garis diagonal dan mengikuti arah diagonal, maka dapat disimpulkan bahwa model regresi berdistribusi normal dan memenuhi asumsi normalitas.

- Pengujian dengan statistik (Kolmogorov-Smirnov)

H0: Error berdistribusi normal

H1: Error tidak berdistribusi normal

p-value (0,200) > alpha (0,05)

Kesimpulan: Gagal Tolak H0, sehingga dapat dinyatakan bahwa biaya yang dikeluarkan untuk pembelian kuota internet berdistribusi normal.

c) Uji Multikolinearitas

| Coefficients ^a | | | | | | | | |
|---------------------------|-----------------|-----------------------------|------------|---------------------------|--------|------|-------------------------|-------|
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
| | | B | Std. Error | Beta | | | Tolerance | VIF |
| 1 | (Constant) | 48616.569 | 6173.243 | | 7.875 | .000 | | |
| | Waktu_Streaming | 5614.814 | 1055.181 | .424 | 5.321 | .000 | .587 | 1.705 |
| | Kuota_Seluler | 717.426 | 154.986 | .388 | 4.629 | .000 | .532 | 1.880 |
| | Kuota_Wifi | -302.145 | 88.061 | -.226 | -3.431 | .001 | .858 | 1.166 |
| | Provider=XL | -12534.158 | 5484.734 | -.142 | -2.285 | .025 | .963 | 1.038 |

a. Dependent Variable: Biaya_Kuota

Tabel 4.2.3. b) Nilai VIF

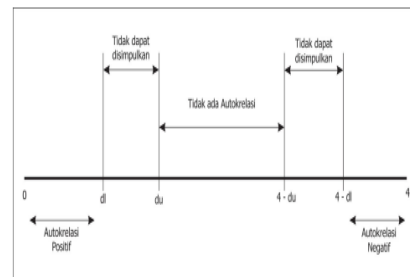
Sesuai dengan syarat Uji Multikolinearitas, jika nilai VIF kurang dari 10 maka dapat disimpulkan bahwa tidak terdapat masalah multikolinearitas.

d) Uji Autokorelasi

| Model Summary ^b | | | | | |
|----------------------------|-------------------|----------|-------------------|----------------------------|---------------|
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | .829 ^a | .687 | .672 | 22228.837 | 2.308 |

a. Predictors: (Constant), Provider=XL, Waktu_Streaming, Kuota_Wifi, Kuota_Seluler

b. Dependent Variable: Biaya_Kuota



Tabel 4.2.3. c) Nilai Durbin-Watson

Gambar 4.2.3. c) Visual Durbin-Watson

$$dL = 1.5627 \quad dU = 1.7501 \quad d = 2.308$$

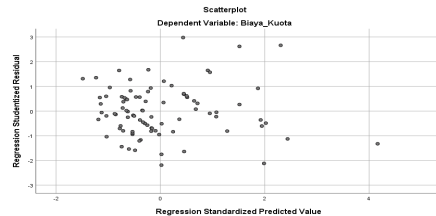
$$4-dL = 2.4373 \quad 4-dU = 2.2499 \quad 4-d = 1.692$$

Interpretasi : Dari hasil statistik didapatkan $4-dU(2.2499) < d(2.308) < 4-dL(2.4373)$ sehingga sesuai dengan tabel uji Durbin-Watson, maka pengujian tidak menghasilkan kesimpulan yang pasti

e) Uji Variansi Error Konstan

| | p-value | Keputusan |
|-----------------|----------|----------------|
| const | 0.000402 | Tolak H0 |
| Waktu_Streaming | 0.197806 | Gagal Tolak H0 |
| Kuota_Wifi | 0.952467 | Gagal Tolak H0 |
| Kuota_Seluler | 0.592760 | Gagal Tolak H0 |
| Provider_XL | 0.237062 | Gagal Tolak H0 |

Tabel 4.2.3 d) uji glejser



Gambar 4.2.3 d) Scatterplot

H0 : Variansi Error Konstan (homoskedastisitas)

H1 : Variansi Error Tidak Konstan (heteroskedastisitas)

- Berdasarkan output di atas diketahui 4 variabel independen memiliki nilai p-value lebih besar dari nilai confidence level (0.05), maka dapat disimpulkan gagal tolak H0. Artinya, tidak ada masalah heteroskedastisitas.
- Dengan melihat grafik scatterplot di atas, titik-titik yang terbentuk menyebar secara acak, baik di atas maupun di bawah angka 0 pada sumbu-Y. Maka dapat disimpulkan bahwa tidak terjadi heteroskedastisitas pada model regresi.

4.2.4 Seleksi Variabel Independen

Dalam melakukan proses variabel selection atau pemilihan variabel, dilakukan menggunakan metode *backward selection*. Maka didapatkan hasil dengan variabel independen signifikan lebih banyak daripada awal regresi.

| Variables Entered/Removed ^a | | | |
|----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|-------------------------------------------------------|
| Model | Variables Entered | Variables Removed | Method |
| 1 | Jaringan=Wifi, Provider=Telkomsel, Provider=Smartfren, Media_Streaming, Kuota_Wifi, Waktu_Medsos, Media_Sosial, Provider=XL, Kuota_Seluler, Waktu_Streaming | . | Enter |
| 2 | . | Media_Sosial | Backward criterion: Probability of F-to-remove = .100 |
| 3 | . | Provider=Smartfren | Backward criterion: Probability of F-to-remove = .100 |
| 4 | . | Media_Streaming | Backward criterion: Probability of F-to-remove = .100 |
| 5 | . | Jaringan=Wifi | Backward criterion: Probability of F-to-remove = .100 |
| 6 | . | Provider=Telkomsel | Backward criterion: Probability of F-to-remove = .100 |
| 7 | . | Waktu_Medsos | Backward criterion: Probability of F-to-remove = .100 |

a. Dependent Variable: Biaya_Kuota
b. All requested variables entered.

Tabel 4.2.4 a) Backward Elimination

| Model Summary ^b | | | | |
|----------------------------|-------------------|----------|-------------------|----------------------------|
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .941 ^a | .707 | .670 | 22302.771 |
| 2 | .941 ^b | .707 | .673 | 22184.218 |
| 3 | .940 ^c | .705 | .675 | 22107.990 |
| 4 | .938 ^d | .703 | .677 | 22046.453 |
| 5 | .938 ^e | .699 | .677 | 22064.292 |
| 6 | .937 ^f | .693 | .675 | 22130.414 |
| 7 | .939 ^g | .687 | .672 | 22228.937 |
| | | | | 2.308 |

a. Predictors: (Constant), Jaringan=Wifi, Provider=Telkomsel, Provider=Smartfren, Media_Streaming, Kuota_Wifi, Waktu_Medsos, Media_Sosial, Provider=XL, Kuota_Seluler, Waktu_Streaming
b. Predictors: (Constant), Jaringan=Wifi, Provider=Telkomsel, Provider=Smartfren, Media_Streaming, Kuota_Wifi, Waktu_Medsos, Provider=XL, Kuota_Seluler, Waktu_Streaming
c. Predictors: (Constant), Jaringan=Wifi, Provider=Telkomsel, Media_Streaming, Kuota_Wifi, Waktu_Medsos, Provider=XL, Kuota_Seluler, Waktu_Streaming
d. Predictors: (Constant), Jaringan=Wifi, Provider=Telkomsel, Waktu_Medsos, Provider=XL, Kuota_Seluler, Waktu_Streaming
e. Predictors: (Constant), Provider=Telkomsel, Kuota_Wifi, Waktu_Medsos, Provider=XL, Kuota_Seluler, Waktu_Streaming
f. Predictors: (Constant), Kuota_Wifi, Waktu_Medsos, Provider=XL, Kuota_Seluler, Waktu_Streaming
g. Predictors: (Constant), Kuota_Wifi, Provider=XL, Kuota_Seluler, Waktu_Streaming
h. Dependent Variable: Biaya_Kuota

Tabel 4.2.4 b) Model Regresi Setelah Backward

Berdasarkan tabel diatas terdapat beberapa variabel yang dieliminasi, dan diakhir *backward elimination* tersisa 4 variabel yaitu Waktu_Streaming, Kuota_Wifi, Kuota_Seluler, dan Provider_XL.

4.2.5 Kebaikan Model Regresi Akhir

| Coefficients ^a | | | | | | | | |
|---------------------------|-----------------|-----------------------------|------------|---------------------------|--------|------|-------------------------|-------|
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
| | | B | Std. Error | Beta | | | Tolerance | VIF |
| 1 | (Constant) | 48616.569 | 6173.243 | | 7.875 | .000 | | |
| | Waktu_Streaming | 5614.814 | 1055.181 | .424 | 5.321 | .000 | .587 | 1.705 |
| | Kuota_Seluler | 717.426 | 154.986 | .388 | 4.629 | .000 | .532 | 1.880 |
| | Kuota_Wifi | -302.145 | 88.061 | -.226 | -3.431 | .001 | .858 | 1.166 |
| | Provider=XL | -12534.158 | 5484.734 | -.142 | -2.285 | .025 | .963 | 1.038 |

a. Dependent Variable: Biaya_Kuota

Tabel 4.2.5 a) Persamaan Regresi Akhir

Persamaan regresi dari output diatas adalah:

$$\text{biaya_kuota} = 48616.569 + 5614.814 (\text{waktu_streaming}) + 717.426 (\text{kuota_seluler}) - 302.145 (\text{kuota_wifi}) - 12534.158 (\text{provider_xl})$$

| Model Summary ^b | | | | | |
|----------------------------|-------------------|----------|-------------------|----------------------------|---------------|
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | .829 ^a | .687 | .672 | 22228.837 | 2.308 |

a. Predictors: (Constant), Provider=XL, Waktu_Streaming, Kuota_Wifi, Kuota_Seluler

b. Dependent Variable: Biaya_Kuota

Tabel 4.2.5 b) Model Regresi Akhir

Berikut adalah interpretasi dari output tersebut:

- Model Regresi Akhir memiliki R-Square sebesar 0.687, yang menunjukkan bahwa model mampu menjelaskan sekitar 68.7% dari variasi dalam variabel target (biaya_kuota).
- Model Regresi Akhir memiliki Adjusted R-Square sebesar 0.672, yang menunjukkan bahwa model mampu menjelaskan sekitar 67.2% dari variasi dalam variabel target setelah mempertimbangkan jumlah variabel independen yang digunakan dalam model.

Secara keseluruhan, model regresi akhir menunjukkan peningkatan adjusted r-squared dari model regresi yang sebelumnya yang artinya semakin baik dalam menjelaskan hubungan antara variabel independen dengan variabel dependen.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan dari hasil survei yang didapat dengan mengambil responden dari 89 mahasiswa di prodi Teknologi Sains Data angkatan 2022, didapatkan hasil sebagai berikut :

1. Analisis eksplorasi data.

Dari hasil analisis eksplorasi data yang dilakukan dengan menyajikan hubungan antar variabel didapatkan bahwa variabel respon berdistribusi normal. Berdasarkan nilai korelasi, dapat diketahui bahwa variabel prediktor yang berkorelasi atau berpengaruh terhadap variabel respons (biaya pembelian kuota bulanan) adalah kuota seluler (0.74), waktu streaming (0.7), dan waktu medsos (0.59).

2. Regresi linier

a. Uji Serentak

Berdasarkan perhitungan diatas, dapat kita lihat bahwa nilai p-value ($2.008259e-20$) lebih kecil dari nilai alpha (0.05) yang artinya kita dapat simpulkan Tolak H_0 . Berarti terdapat hubungan antara variabel independen dengan variabel dependen.

b. Uji Parsial

Didapatkan bahwa variabel waktu_streaming memiliki hubungan yang linier dengan biaya pengeluaran dalam pembelian kuota internet bulanan. Sedangkan pada variabel kuota_wi-fi, kuota_seluler, dan provider_xl tidak memiliki hubungan yang linier dengan biaya pengeluaran dalam pembelian kuota internet bulanan.

c. Uji Asumsi Error

- Uji Linearitas berdasarkan visualisasi bahwa variabel waktu_streaming dan kuota_wifi memiliki hubungan yang linier dengan biaya pengeluaran dalam pembelian kuota internet bulanan. Sedangkan pada variabel kuota_seluler tidak memiliki hubungan yang linier dengan biaya pengeluaran dalam pembelian kuota internet bulanan.
- Uji Independensi Error berdasarkan hasil statistik didapatkan $4-dU(2.27) < d(2.31) < 4-dL(2.41)$ sehingga sesuai dengan tabel uji Durbin-Watson, maka pengujian tidak menghasilkan kesimpulan yang pasti
- Uji Variansi Error Konstan dilakukan uji glejser dan berdasarkan visualisasi dapat disimpulkan terjadi

homoskedastisitas antar variabel independen dalam model regresi.

- Uji Normalitas hasil dilakukan pengujian Kolmogorov Smirnov didapatkan p-value (0,200) yang lebih besar dari alpha (0,05). Artinya Gagal Tolak H_0 , sehingga dapat dinyatakan bahwa biaya yang dikeluarkan untuk pembelian kuota internet berdistribusi normal.
 - Uji Multikolinearitas dilihat dengan melihat nilai VIF, sesuai dengan Uji Multikolinearitas, jika nilai VIF kurang dari 10 maka dapat disimpulkan bahwa tidak terdapat masalah multikolinearitas.
- d. Kebaikan model dapat dijelaskan bahwa nilai R-Squared dan Adjusted R-Square sebesar 0.687 dan 0.672, yang menunjukkan bahwa model mampu menjelaskan sekitar 68.7%, sedangkan 31,3% dijelaskan oleh variabel lain yang tidak termasuk dalam perhitungan dari variasi dalam variabel target (biaya_kuota).

Sehingga, dapat dijelaskan secara keseluruhan, model regresi akhir menunjukkan peningkatan adjusted r-squared dari model regresi yang sebelumnya yang artinya semakin baik dalam menjelaskan hubungan antara variabel independen dengan variabel dependen.

5.2 Saran

Dari hasil analisa yang telah dilakukan, untuk pemenuhan kesimpulan pada uji autokorelasi bisa diatasi dengan menggunakan uji *run-test*. Dan kami peneliti menyarankan untuk mengurangi waktu yang digunakan untuk streaming, memilih provider xl, serta penggunaan kuota data seluler dan kuota wi-fi untuk menghemat biaya pengeluaran dalam pembelian kuota internet bulanan.

Kami harap hasil penelitian kami dapat digunakan sebagai referensi untuk penelitian selanjutnya. Kami juga berharap hasil penelitian kami dapat berguna bagi pembaca untuk membuat keputusan dalam manajemen pengeluaran dalam pembelian kuota internet bulanan.

DAFTAR PUSTAKA

- D Agustiah, T Fauzi, E Ramadhani. (2020). Dampak Penggunaan Media Sosial Terhadap Perilaku Belajar Siswa .Jurnal Bimbingan dan Konseling Islam, 4(2), 181-190.
- Montgomery Douglas C., Elizabeth A. Peck, G. Geoffrey Vinning,(2021), *Introduction to Linear Regression Analysis Wiley Series in Probability and Statistics*, John Wiley & Sons, New Jersey, 532, 617.
- Nasrullah, Rulli. (2015). Media Sosial Perspektif Komunikasi, Budaya dan Sosioteknologi. Bandung: Simbiosis Rekatama Media.
- Nugraha, Billy. (2022). Pengembangan Uji Statistik: Implementasi Metode Regresi Linear Berganda dengan Pertimbangan Uji Asumsi Klasik, Pradina Pustaka, Karawang, 8-16.
- Suryadi, Sudi. (2015). Peranan Perkembangan Teknologi Informasi dan Komunikasi Dalam Kegiatan Pembelajaran dan Perkembangan Dunia Pendidikan, Jurnal Universitas Labuan Bajo, 3(3), 12-16.
- Wood, Andrew F., Matthew J. Smith (2004), Online Communication. London: Taylor & Francis.

LAMPIRAN

Codingan dan Dataset:

<https://drive.google.com/drive/folders/1HHKpAi18yzlJYSLv-gk00Owi0-xIkRv8?usp=sharing>

Kuisisioner :

https://docs.google.com/forms/d/e/1FAIpQLSfqGBZKJtXFnyUQ3-sktv1H_Lu2suGldhvOjmqMWjMEsOxmQ/viewform