

---

**OPTIMASI ANALISIS KETERLAMBATAN PENERBANGAN PADA DATASET  
FLIGHT DELAYS MELALUI DASHBOARD DATA ANALYTICS BERBASIS  
LOOKER STUDIO**

Wina Elsa Wardana<sup>1</sup>

<sup>1</sup>Universitas Merdeka Malang

Email: [winaelsaaw18@gmail.com](mailto:winaelsaaw18@gmail.com)

**Abstrak:** Perkembangan teknologi digital di era Revolusi Industri 4.0 telah mendorong pemanfaatan data secara intensif di berbagai sektor, termasuk industri penerbangan. Salah satu masalah utama yang dihadapi maskapai penerbangan adalah keterlambatan penerbangan, yang berdampak signifikan pada operasional, biaya, dan kepuasan pelanggan. Studi ini bertujuan untuk menganalisis pola keterlambatan penerbangan menggunakan dataset Keterlambatan Penerbangan dari Kaggle dan merancang dashboard interaktif berbasis Looker Studio untuk mendukung pengambilan keputusan. Metode yang digunakan meliputi pengumpulan data, pembersihan data, analisis statistik (One Way ANOVA, regresi linier, dan uji Chi-Square), dan visualisasi data. Hasil menunjukkan bahwa terdapat perbedaan signifikan dalam rata-rata keterlambatan antar maskapai penerbangan, jarak penerbangan memiliki pengaruh yang sangat lemah terhadap keterlambatan, dan bandara asal memiliki hubungan yang signifikan dengan terjadinya keterlambatan. Dashboard yang dihasilkan mampu memberikan wawasan komprehensif mengenai faktor-faktor keterlambatan sehingga dapat digunakan dalam mengoptimalkan operasional penerbangan.

**Kata Kunci:** Data Analytics, Flight Delays, One Way ANOVA, Regresi Linear, Chi-Square.

***Abstract:** The development of digital technology in the Industrial Revolution 4.0 era has driven the intensive use of data in various sectors, including the aviation industry. One of the main problems faced by airlines is flight delays, which have a significant impact on operations, costs, and customer satisfaction. This study aims to analyze flight delay patterns using the Flight Delay dataset from Kaggle and design an interactive dashboard based on Looker Studio to support decision-making. The methods used include data collection, data cleaning, statistical analysis (One Way ANOVA, linear regression, and Chi-Square test), and data visualization. The results show that there are significant differences in the average delay between airlines, flight distance has a very weak effect on delays, and the airport of origin has a significant relationship with the occurrence of delays. The resulting dashboard is able to provide comprehensive insights into delay factors so that it can be used to optimize flight operations.*

***Keywords:** Data Analytics, Flight Delays, One Way ANOVA, Linear Regression, Chi-Square.*

## PENDAHULUAN

Perkembangan teknologi digital pada era Revolusi Industri 4.0 telah mendorong transformasi besar dalam berbagai sektor, termasuk industri penerbangan. Integrasi data, teknologi data, teknologi informasi, dan sistem otomatisasi kini menjadi fondasi penting dalam meningkatkan efisiensi operasional serta kualitas layanan. Dalam konteks ini, kemampuan mengolah dan menganalisis data menjadi kompetensi utama yang dibutuhkan organisasi untuk meningkatkan daya saing dan mendukung proses pengambilan keputusan berbasis bukti (data-driven decision making).

Salah satu tantangan signifikan dalam operasional maskapai adalah keterlambatan penerbangan (flight delays). Keterlambatan ini tidak hanya berdampak pada kenyamanan dan kepuasan penumpang, tetapi juga menimbulkan konsekuensi operasional seperti penumpukan jadwal, peningkatan biaya bahan bakar, gangguan rotasi pesawat, hingga kerugian finansial. Dengan memanfaatkan analisis data, maskapai dapat memahami pola keterlambatan, mengidentifikasi faktor-faktor penyebab, serta memprediksi potensi delay untuk meningkatkan efektivitas perencanaan dan mitigasi gangguan operasional.

Dalam beberapa penelitian terkini, penerapan metode machine learning seperti deep learning terbukti efektif dalam memprediksi keterlambatan penerbangan dengan mempertimbangkan berbagai faktor operasional dan lingkungan secara komprehensif.

Penelitian ini dilaksanakan melalui beberapa tahapan data sains diantaranya pengumpulan data melalui web terbuka (Kaggle) dataset, melakukan pra-proses data dan analisis data dengan menggunakan tools Google Colaboratory, dan visualisasi data menggunakan tools Looker Studio. Analisis data yang dilakukan meliputi analisis One-Way ANOVA, Uji Regresi, dan Chi-Square.

Hasil penelitian ini diharapkan memiliki manfaat bagi kebutuhan operasional maskapai secara luas, dimana rekomendasi strategi keputusan operasional maskapai penerbangan melalui aspek-aspek yang dianalisis.

## METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif yang merupakan pendekatan penelitian yang mewakili paham positivisme dan alat uji statistik, maupun matematik yang sering disebut sebagai analisis deskriptif kuantitatif.

## Sumber Data

Penelitian ini menggunakan dataset yang bersumber dari platform Kaggle. Dataset Flight Delays merupakan sekumpulan data operasional penerbangan yang berisi informasi mengenai jadwal keberangkatan, kedatangan, maskapai, bandara asal, dan tujuan, serta berbagai indikator keterlambatan.

## Tahap Pengolahan Data

Pada tahap ini, dilakukan pemeriksaan dan pembersihan data agar tidak terjadi inkonsistensi data pada saat dilakukan analisis lanjutan. Dengan melakukan pemeriksaan missing value, outlier, pembersihan duplikasi data dan penanganannya.

## Teknik Analisis Statistik

Penelitian ini menggunakan tiga Teknik analisis statistik yaitu, *One-Way* ANOVA, Regresi Linear, dan uji *Chi-Square*. *One-Way* ANOVA merupakan metode statistik yang digunakan untuk menguji perbedaan rata-rata lebih dari dua kelompok [5]. Pada penelitian ini *One-Way* ANOVA digunakan untuk mengetahui perbedaan rata-rata keterlambatan antar maskapai. Regresi Linear adalah suatu metode analisis statistik yang mempelajari pola hubungan antara dua atau lebih variabel [6]. Pada penelitian ini, Regresi Linear digunakan untuk memeriksa pengaruh jarak penerbangan dengan keterlambatan. Sedangkan, uji *Chi-Square* adalah salah satu jenis uji komparatif non parametris yang dilakukan pada dua variabel, dimana skala data kedua variabel adalah nominal. Pada penelitian ini, uji *Chi-Square* digunakan untuk menguji hubungan antara bandara asal dan keterlambatan keberangkatan.

## Visualisasi Dashboard

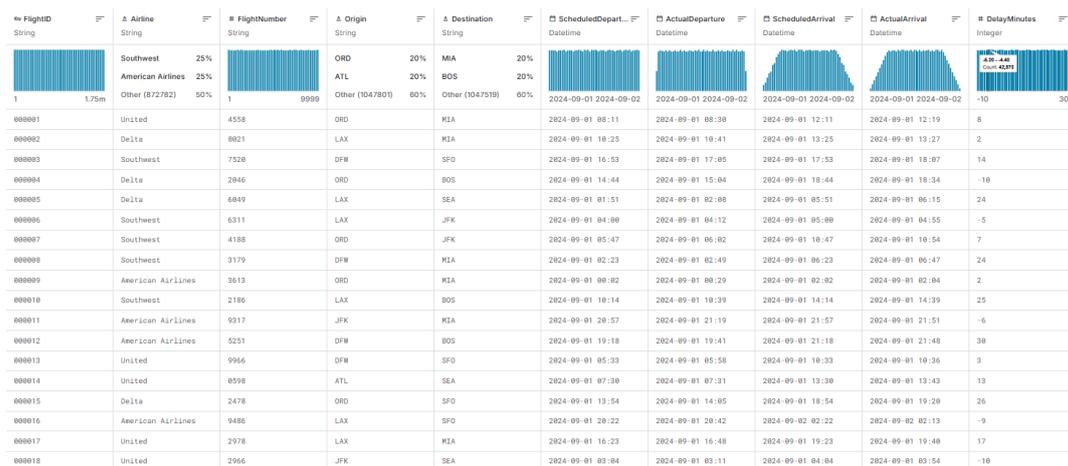
Dashboard adalah satu aplikasi yang secara real time akan memonitoring berbagai informasi yang dibutuhkan oleh suatu organisasi atau perusahaan dengan berbagai macam format. Pada penelitian ini dashboard dirancang menggunakan Looker Studio untuk menampilkan insight mengenai total penerbangan, rata-rata keterlambatan, bandara dengan delay tertinggi, dan performa maskapai.

## HASIL DAN PEMBAHASAN

Dalam melakukan analisis memerlukan beberapa langkah. Langkah pertama adalah pengumpulan data, yang merupakan dasar utama penelitian karena kualitas dan keakuratannya

memengaruhi validitas dan keakuratan temuan. Data yang digunakan dalam analisis penelitian ini adalah dataset yang diperoleh dari platform Kaggle berjudul "Flight Delays."

Dataset "Flight Delays" adalah kumpulan data operasional penerbangan yang berisi informasi tentang jadwal keberangkatan dan kedatangan, maskapai penerbangan, bandara asal dan tujuan, dan berbagai indikator keterlambatan.



FlightID	Airline	FlightNumber	Origin	Destination	ScheduledDeparture	ActualDeparture	ScheduledArrival	ActualArrival	DelayMinutes
000001	United	4558	ORD	MIA	2024-09-01 08:11	2024-09-01 08:38	2024-09-01 12:11	2024-09-01 12:19	8
000002	Delta	8821	LAX	MIA	2024-09-01 18:25	2024-09-01 18:41	2024-09-01 13:25	2024-09-01 13:27	2
000003	Southwest	7528	DFW	SFO	2024-09-01 16:53	2024-09-01 17:05	2024-09-01 17:53	2024-09-01 18:07	14
000004	Delta	2946	ORD	BOS	2024-09-01 14:44	2024-09-01 15:04	2024-09-01 18:44	2024-09-01 18:34	-10
000005	Delta	6849	LAX	SEA	2024-09-01 01:51	2024-09-01 02:08	2024-09-01 05:51	2024-09-01 06:15	24
000006	Southwest	6311	LAX	JFK	2024-09-01 04:08	2024-09-01 04:12	2024-09-01 05:08	2024-09-01 04:55	-5
000007	Southwest	4188	ORD	JFK	2024-09-01 05:47	2024-09-01 06:02	2024-09-01 10:47	2024-09-01 10:54	7
000008	Southwest	3179	DFW	MIA	2024-09-01 02:23	2024-09-01 02:49	2024-09-01 06:23	2024-09-01 06:47	24
000009	American Airlines	3613	ORD	MIA	2024-09-01 08:02	2024-09-01 08:29	2024-09-01 02:02	2024-09-01 02:04	2
000010	Southwest	2186	LAX	BOS	2024-09-01 18:14	2024-09-01 18:39	2024-09-01 14:14	2024-09-01 14:39	25
000011	American Airlines	9317	JFK	MIA	2024-09-01 20:57	2024-09-01 21:19	2024-09-01 21:57	2024-09-01 21:51	-6
000012	American Airlines	5251	DFW	BOS	2024-09-01 19:18	2024-09-01 19:41	2024-09-01 21:18	2024-09-01 21:48	30
000013	United	9966	DFW	SFO	2024-09-01 05:33	2024-09-01 05:58	2024-09-01 10:33	2024-09-01 10:36	3
000014	United	8598	ATL	SEA	2024-09-01 07:38	2024-09-01 07:31	2024-09-01 13:38	2024-09-01 13:43	13
000015	Delta	2478	ORD	SFO	2024-09-01 13:54	2024-09-01 14:05	2024-09-01 18:54	2024-09-01 19:28	26
000016	American Airlines	9406	LAX	SFO	2024-09-01 20:22	2024-09-01 20:42	2024-09-02 02:22	2024-09-02 02:13	-9
000017	United	2978	LAX	MIA	2024-09-01 16:23	2024-09-01 16:48	2024-09-01 19:23	2024-09-01 19:40	17
000018	United	2966	JFK	SEA	2024-09-01 03:04	2024-09-01 03:11	2024-09-01 04:04	2024-09-01 03:54	-10

Gambar 1. Dataset *Flight Delays by Kaggle*

Setelah tahap pengumpulan data selesai, langkah selanjutnya adalah mengembangkan pertanyaan analisis dan menentukan tujuan utama analisis. Langkah selanjutnya adalah pembersihan data. Proses ini mengatasi nilai yang hilang dan outlier, yang dapat memengaruhi kualitas analisis dan validitas hasil penelitian. Nilai yang hilang merupakan masalah utama dalam pemrosesan data karena dapat menyebabkan bias dalam interpretasi dan mengurangi akurasi model prediktif yang digunakan dalam analisis lebih lanjut.

Sebelum pembersihan, dataset yang digunakan dalam penelitian ini berisi 468.873 nilai yang hilang pada kolom alasan keterlambatan. Oleh karena itu, pembersihan data dilakukan untuk memastikan data siap digunakan dengan memeriksa tipe data di setiap kolom untuk kesesuaian dan memeriksa outlier untuk mencegah inkonsistensi data. Pada tahap pertama pembersihan data, dilakukan pengecekan data. Identifikasi kolom data digunakan untuk menentukan masalah yang mendasarinya. Berikut adalah hasil eksplorasi awal menggunakan fungsi `df.info()`. Dataset Keterlambatan Penerbangan terdiri dari 265.542 baris dan 16 kolom.

Setelah memeriksa informasi di setiap kolom dataset penundaan penerbangan, langkah selanjutnya dalam pembersihan data adalah memeriksa nilai yang hilang. Nilai yang hilang diperiksa menggunakan Google Colab dengan kode `df.isnull().sum()`. Metode ini digunakan

untuk menentukan jumlah nilai yang hilang di setiap kolom dataset *flight delays*. Setelah diperiksa, ditemukan 468.873 nilai yang hilang di kolom DelayReason.

Karena ada nilai yang hilang di kolom DelayReason, untuk menghindari inkonsistensi data selama analisis, nilai yang hilang harus ditangani. Hal ini dilakukan dengan mengisi nilai yang hilang dengan nilai lain. Kemudian, periksa kembali apakah nilai yang hilang telah ditangani dengan benar dengan memeriksa menggunakan kode `df.isnull().sum()`. Pastikan bahwa setiap kolom dalam dataset penundaan penerbangan tidak lagi mengandung nilai yang hilang dan kemudian identifikasi konsistensi data.

Setelah tahap pembersihan data selesai, data siap diproses untuk menghasilkan informasi yang siap untuk dianalisis. Tahapan pengolahan data menggunakan teknik statistik deskriptif, pengelompokan data, pembuatan variabel baru, dan penyusunan kumpulan data untuk analisis lebih lanjut. Berikut adalah ringkasan langkah-langkah pengolahan data :

## A. Statistik Deskriptif Kumpulan Data

Kumpulan data yang telah dibersihkan dianalisis menggunakan statistik deskriptif untuk menentukan nilai rata-rata, minimum, dan maksimum, deviasi standar, dan distribusi nilai untuk variabel numerik seperti DelayMinutes, Distance, ArrivalDelay, dan DepartureDelay. Tahapan ini memberikan gambaran umum tentang karakteristik data setelah proses pembersihan.

## B. Pengelompokan Data

Untuk menjawab perumusan masalah penelitian, beberapa pengelompokan dilakukan sebagai berikut:

### 1. Rata-rata keterlambatan kedatangan per maskapai penerbangan

Data dikelompokkan berdasarkan nama maskapai penerbangan untuk menentukan maskapai penerbangan mana yang memiliki rata-rata keterlambatan tertinggi. Hasil pengolahan ini digunakan untuk visualisasi grafik batang.

### 2. Jumlah penerbangan yang tertunda per bandara asal

Sebuah flag dibuat untuk menunjukkan apakah suatu penerbangan tertunda atau tidak. Setelah itu, data dikelompokkan berdasarkan bandara asal untuk menghitung jumlah penerbangan yang mengalami keterlambatan keberangkatan di setiap bandara.

### 3. Pengelompokan Jarak Penerbangan

Jarak penerbangan dibagi menjadi beberapa kategori (pendek, menengah, panjang, dan sangat panjang). Setelah pengelompokan, durasi keterlambatan rata-rata untuk setiap kelompok jarak dihitung.

Tujuannya adalah untuk menentukan apakah jarak penerbangan memengaruhi tingkat keterlambatan.

### C. Analisis Hubungan dan Korelasi Antar Variabel

Perhitungan korelasi dilakukan untuk memeriksa hubungan antara jarak penerbangan dan keterlambatan kedatangan. Selanjutnya, visualisasi scatterplot dengan garis tren regresi dibuat untuk menunjukkan bagaimana perubahan jarak berhubungan dengan perubahan tingkat keterlambatan. Tahap ini berfungsi sebagai bagian dari pemrosesan sebelum memasuki uji regresi linier.

### D. Persiapan Data untuk Uji Statistik

Beberapa langkah disiapkan untuk analisis lebih lanjut, seperti:

#### 1) Persiapan Data untuk Uji ANOVA

Dataset diatur ulang menjadi dua kolom (Maskapai Penerbangan dan Keterlambatan Kedatangan) untuk digunakan dalam uji ANOVA, yang bertujuan untuk menentukan apakah ada perbedaan keterlambatan rata-rata antar maskapai penerbangan.

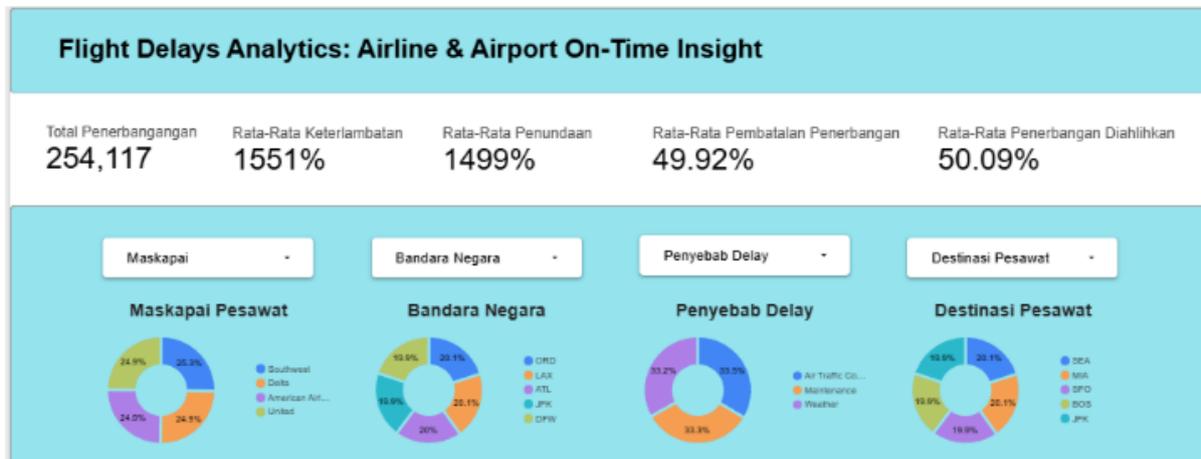
#### 2) Persiapan Data untuk Uji Korelasi dan Regresi

Data jarak dan keterlambatan kedatangan difilter agar hanya berisi nilai yang valid. Variabel-variabel ini kemudian dipersiapkan untuk analisis regresi linier.

#### 3) Persiapan Data untuk Uji Chi-Square

Variabel kategorikal tambahan dibuat untuk menunjukkan apakah penerbangan mengalami keterlambatan atau tidak. Variabel ini kemudian dipasangkan dengan variabel maskapai penerbangan untuk membentuk tabel kontingensi yang digunakan dalam uji Chi-Square.

Setelah semua pemrosesan selesai, dataset yang telah dibersihkan dan diproses disimpan sebagai file CSV. Dataset ini digunakan untuk visualisasi dasbor di Looker Studio.



Gambar 2. Dashboard *Flight Delays Analytics*

Setelah visualisasi, beberapa hasil analisis statistik diperoleh dari dataset penundaan penerbangan yang menjawab pertanyaan analisis yang telah ditentukan, yang diuraikan sebagai berikut :

1. Pertanyaan Analisis 1

Maskapai mana yang memiliki rata-rata keterlambatan kedatangan tertinggi?

Analisis dengan Uji ANOVA (One-Way ANOVA) dan didapatkan hipotesis sebagai berikut :

- H<sub>0</sub> : Tidak ada perbedaan signifikan rata-rata ArrivalDelay antar maskapai (Semua maskapai memiliki rata-rata keterlambatan kedatangan yang sama.).
- H<sub>1</sub>: Ada perbedaan signifikan rata-rata ArrivalDelay antar maskapai.

```
Data Loaded.
Airline origin destination delayminutes delayreason cancelled \
0 United ORD MIA 8.00 Weather True
1 Delta LAX MIA 2.00 Air Traffic Control True
2 Southwest DFW SFO 14.00 Weather True
3 Delta LAX SFA 24.00 Air traffic control false
4 Southwest ORD JFK 7.00 Weather False

Diverted Distance DepartureDelay ArrivalDelay
0 False 1831.00 19.00 8.00
1 True 1886.00 16.00 2.00
2 True 2588.00 12.00 14.00
3 True 2258.00 17.00 24.00
4 True 1674.00 15.00 7.00

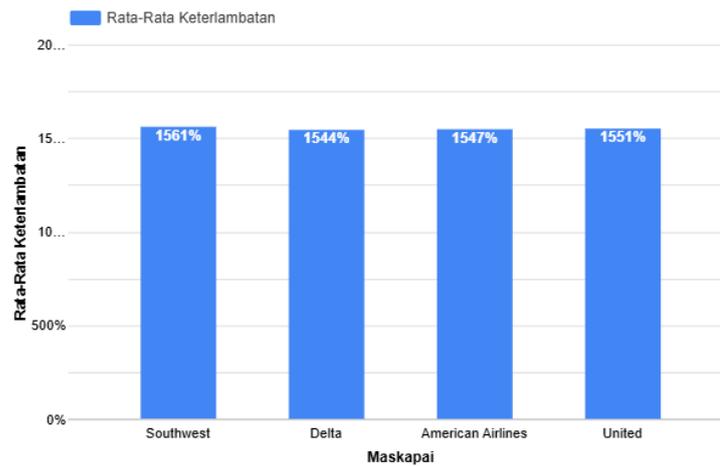
Kolom dataset: ['Airline', 'Origin', 'Destination', 'DelayMinutes', 'DelayReason', 'Cancelled', 'Diverted', 'Distance', 'DepartureDelay', 'ArrivalDelay']

=== One-way ANOVA: ArrivalDelay - Airline ===
SUM_Sq df F PR(>F)
C(Airline) 523.81 3.00 2.34 0.87
Residual 14511532.79 194621.00 NaN NaN

=== Keputusan Hipotesis ===
+ Gagal tolak H0: Tidak ada perbedaan signifikan antar maskapai.
/tmp/ipython-input-734286138.py:23: FutureWarning: Series.__getitem__ treating keys as positions is deprecated. In a future version, integer keys will all
p_value = anova_table["PR(>F)"][0]
```

Gambar 3. One Way ANOVA

Dasar pengujian pada Google Colab yang menghasilkan p-value dari uji ANOVA PR(>F) = 0.000000 . Maka dari itu, H0 ditolak dan H1 diterima karena p-value yang dihasilkan < 0.05. Sehingga dapat disimpulkan adanya perbedaan signifikan rata-rata keterlambatan kedatangan antar maskapai.



Gambar 4. Bar Chart Rata-Rata Delays Per Maskapai

Berdasarkan grafik diatas, Southwest memiliki rata-rata keterlambatan kedatangan (ArrivalDelay) paling tinggi dan merupakan performa ketepatan waktu yang paling buruk.

#### E. Pertanyaan Analisis 2

Apakah jarak penerbangan berpengaruh terhadap kedatangan?

Analisis dengan Korelasi Pearson + Regresi Linear dan didapatkan hipotesis sebagai berikut :

- H0: Distance tidak berpengaruh terhadap ArrivalDelay
- H1: Distance berpengaruh terhadap ArrivalDelay

```
*** == Korelasi Pearson: Distance vs ArrivalDelay ==
r = -0.0012
p-value = 0.593418
+ Gagal tolak H0: Tidak ada hubungan signifikan.

== Hasil Regresi Linear ==
                                OLS Regression Results
-----
Dep. Variable:      ArrivalDelay    R-squared:      0.000
Model:              OLS             Adj. R-squared: -0.000
Method:             Least Squares   F-statistic:    0.2850
Date:               Mon, 17 Nov 2025 Prob (F-statistic): 0.593
Time:               05:36:42        Log-Likelihood: -6.9589e+05
No. Observations:  194625          AIC:            1.392e+06
Df Residuals:      194623          BIC:            1.392e+06
Df Model:           1
Covariance Type:   nonrobust

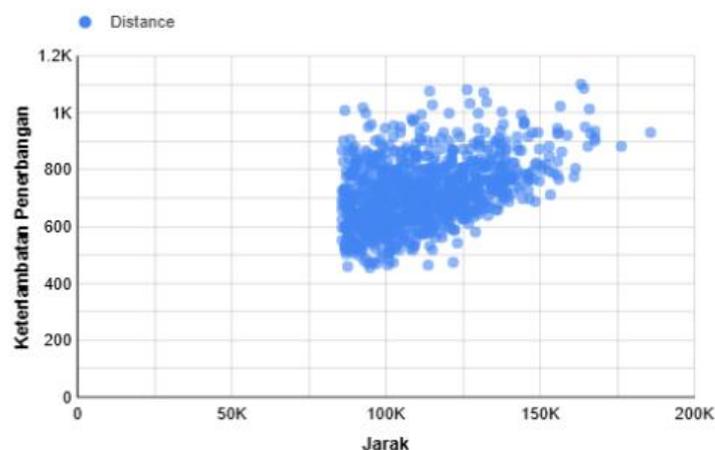
=====
                coef      std err      t      P>|t|      [0.025      0.975]
-----
const          15.5380      0.041    376.789   0.000    15.457    15.619
Distance      -1.249e-05      2.34e-05   -0.534   0.593   -5.83e-05  3.34e-05
=====

Omnibus:          174050.930    Durbin-Watson:    2.003
Prob(Omnibus):    0.000    Jarque-Bera (JB): 11693.464
Skew:             -0.001    Prob(JB):         0.00
Kurtosis:         1.799    Cond. No.         3.71e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 3.71e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Gambar 5. Regresi Linear

Dasar pengujian pada Google Colab menghasilkan korelasi yang lemah dengan p-value = 0.000000 dengan regresi linear sangat kecil (mendekati 0). Maka dari itu, H<sub>0</sub> ditolak dan H<sub>1</sub> diterima karena p-value < 0.05. Dapat disimpulkan bahwa ada pengaruh signifikan secara statistik antara jarak penerbangan dan keterlambatan kedatangan. Namun nilai R<sup>2</sup> (regresi linear) yang sangat kecil menunjukkan bahwa pengaruhnya sangat lemah.



Gambar 6. Scatterplot Korelasi Jarak dan Delays

Berdasarkan dari grafik scatterplot menunjukkan titik-titik menyebar luas Titik-titik tidak membentuk pola garis yang jelas. Banyak penyebaran ke atas, bawah, kiri, kanan. Tidak ada trend yang kuat. Ini menunjukkan hubungan walaupun ada, itu sangat lemah. Walaupun memang ada hubungan karena garis tren naik sedikit Hal ini mendukung pernyataan bahwa “Jarak memang berpengaruh, tetapi pengaruhnya lemah atau bukan faktor utama penyebab keterlambatan”.

#### F. Pertanyaan Analisis 3

Bandara mana yang paling sering mengalami penundaan keberangkatan?

Analisis dengan uji Chi-Square dan didapatkan hipotesis sebagai berikut :

- H0 : Tidak ada hubungan antara bandara asal (Origin) dan kejadian delay.
- H1 : Ada hubungan antara bandara asal dan kejadian delay.

```
*** Data Loaded for Chi-Square
=== Tabel kontingensi (Airline vs DelayFlag) ===
DelayFlag      1
Airline
American Airlines  48482
Delta              48511
Southwest         48929
United            48783

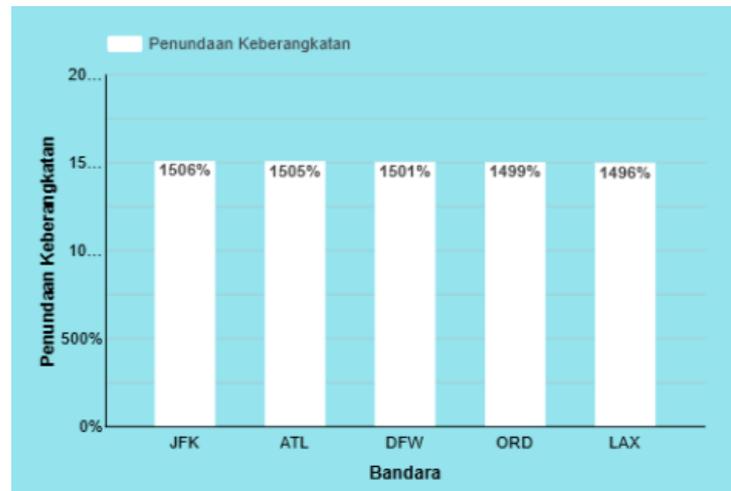
=== Hasil Chi-Square Test ===
Chi-square Statistic : 0.0000
Degrees of Freedom   : 0
P-Value              : 1.000000

=== Keputusan Hipotesis ===
+ Gagal tolak H0: Tidak ada hubungan signifikan antara maskapai dan keterlambatan.

=== Expected Frequency Table ===
DelayFlag      1
Airline
American Airlines 48482.00
Delta             48511.00
Southwest         48929.00
United            48783.00
```

Gambar 7. Uji Chi-Square

Dasar pengujian pada Google Colab menghasilkan p-value = 0.000000. Maka dari itu, H0 ditolak dan H1 diterima karena p-value < 0.05. . Dapat disimpulkan bahwa ada hubungan signifikan antara bandara asal dan jumlah penerbangan yang mengalami keterlambatan keberangkatan.



Gambar 8. Bar Chart Bandara dengan Delay Tertinggi

Berdasarkan grafik diatas dapat disimpulkan bahwa Bandara JFK (John F. Kennedy International Airport) memiliki presentase rata-rata delay tertinggi dan merupakan performa paling buruk dibanding bandara lainnya.

## KESIMPULAN

Southwest menunjukkan kinerja ketepatan waktu terburuk dibandingkan dengan maskapai lain, yang mengindikasikan inefisiensi operasional, kemungkinan dampak keterlambatan pesawat, dan beban rute yang tinggi, yang mengakibatkan waktu putar balik yang lebih singkat dan meningkatkan potensi keterlambatan.

Sementara itu, jarak bukanlah faktor utama dalam keterlambatan, karena keterlambatan lebih sering dipengaruhi oleh kondisi cuaca di bandara keberangkatan dan kedatangan, faktor operasional maskapai, dan manajemen lalu lintas udara atau kemacetan bandara, yang dapat menunda lepas landas dan pendaratan.

JFK adalah bandara tersibuk di antara perbandingan dan oleh karena itu lebih rentan terhadap kemacetan. Volume penerbangan yang tinggi dikombinasikan dengan cuaca ekstrem yang sering terjadi di wilayah tersebut merupakan faktor kunci yang berkontribusi pada tingginya jumlah keterlambatan. Situasi ini semakin diperparah oleh ketergantungan pada infrastruktur yang relatif tua dan lalu lintas internasional yang tinggi, yang menambah kompleksitas operasional bandara.

**DAFTAR PUSTAKA**

- R. A. Valdes, V. Fernando, and G. Comendador, “AVIATION 4 . 0 : MORE SAFETY THROUGH AUTOMATION AND DIGITIZATION,” vol. 174, pp. 225–236, 2018, doi: 10.2495/SAFE170211.
- J. R. Informatika, “PREDICTION OF FLIGHT DELAYS USING FEATURE ENGINEERING , CATBOOST , AND BAYESIAN OPTIMIZATION TO IMPROVE MODEL,” vol. 7, no. 2, pp. 1–8, 2025.
- K. Przanowski and D. Deczyjnych, “PREDICTIVE MODELING OF FLIGHT DELAYS USING DECISSION TREE,” vol. 4, no. 4, pp. 0–3, 2023, doi: 10.55676/asi.v4i2.79.
- U. Negeri, S. Thaha, S. Jambi, and P. Kuantitatif-kualitatif, “Pendekatan penelitian kuantitatif dan kualitatif serta tahapan penelitian,” vol. 15, no. 1, pp. 82–92, 2024.
- Y. Nainggolan, D. L. Hutapea, W. F. Sirait, and M. Sirait, “Anava Satu Jalur ( One Way – Anova ),” vol. 5, pp. 5670–5682, 2025.
- D. P. Egistin, M. Y. Rauza, R. H. Ramadhan, S. Ramadani, and K. Kunci, “Analisis regresi linier sederhana dan penerapannya,” vol. 1, no. 2, pp. 69–78, 2025.
- J. H. Page, “Jurnal Penelitian Sains,” vol. 25, no. April, pp. 120–124, 2023.
- I. Handayani, D. Kurniati, and S. A. Yakin, “Penerapan Dashboard Sebagai Media Pengumpulan Data Laporan Penulisan Mahasiswa di Widuri Menggunakan Rinfo Spreadsheet Pada Perguruan Tinggi,” vol. 2, no. 2, pp. 107–118, 2017.
- “Jurnal Edu Research Indonesian Institute For Corporate Learning And Studies (IICLS) Page 110,” vol. 5, no. September, pp. 110–116, 2024.
- J. Elektronika and D. A. N. Komputer, “Mengoptimalkan Proses Pembersihan Data dalam Analisis Big Data Menggunakan Pipeline Berbasis AI,” vol. 17, no. 2, pp. 657–666, 2024.