

Analisis Statistik Uji Normalitas dan Homogenitas Data Nilai Mata Pelajaran dengan Menggunakan Python

Andy Agustian¹, Kania Lisdiana², Adang Suryana³, Muhammad Nursalman⁴

Universitas Pendidikan Indonesia, Bandung, Indonesia

andyagustian@upi.edu¹, kanialisdi@upi.edu², adangsurjana@upi.edu³, mnursalman@upi.edu⁴

Abstract

Statistical testing, particularly homogeneity and normality testing, is an important step in the data analysis process to ensure the validity of statistical assumptions before conducting further analysis. This article discusses the implementation of Python-based homogeneity and normality testing, a programming language that is increasingly popular among researchers and statistical practitioners due to its ability to handle data analysis efficiently. This research utilizes Python libraries, such as SciPy, NumPy, and Matplotlib, to perform normality testing using the Shapiro-Wilk method, as well as homogeneity testing using the Levene test. These procedures were implemented through Python scripts designed to automate the data analysis process, reduce the potential for manual errors, and improve the reproducibility of results. The results show that the Python-based approach provides flexibility and high accuracy in data analysis, with the ability to handle large and complex datasets. Case studies on real datasets reveal that using Python can accelerate the statistical evaluation process, from data exploration to decision-making based on statistical test results. This article is expected to be a practical guide for researchers and statistical practitioners in utilizing Python to improve the efficiency and reliability of data analysis, especially in homogeneity and normality testing.

Keywords: Statistical Testing, Python, Homogeneity, Normality, Data Analysis

Abstrak

Pengujian statistik, khususnya pengujian homogenitas dan normalitas, merupakan langkah penting dalam proses analisis data untuk memastikan validitas asumsi statistik sebelum melakukan analisis lebih lanjut. Artikel ini membahas implementasi pengujian homogenitas dan normalitas berbasis Python, sebuah bahasa pemrograman yang semakin populer di kalangan peneliti dan praktisi statistik karena kemampuannya dalam menangani analisis data secara efisien. Penelitian ini memanfaatkan pustaka Python, seperti SciPy, NumPy, dan Matplotlib, untuk melakukan pengujian normalitas menggunakan metode Shapiro-Wilk, serta pengujian homogenitas menggunakan uji Levene. Prosedur ini diimplementasikan melalui skrip Python yang dirancang untuk mengotomasi proses analisis data, mengurangi potensi kesalahan manual, dan meningkatkan reproducibility hasil. Hasil penelitian menunjukkan bahwa pendekatan berbasis Python memberikan fleksibilitas dan akurasi tinggi dalam analisis data, dengan kemampuan untuk menangani dataset berukuran besar dan kompleks. Studi kasus pada dataset nyata mengungkapkan bahwa penggunaan Python dapat mempercepat proses evaluasi statistik, mulai dari eksplorasi data hingga pengambilan keputusan berbasis hasil uji statistik. Artikel ini diharapkan dapat menjadi panduan praktis bagi peneliti dan praktisi statistik dalam memanfaatkan Python untuk meningkatkan efisiensi dan keandalan analisis data, khususnya dalam pengujian homogenitas dan normalitas.

Kata kunci: Pengujian Statistik, Python, Homogenitas, Normalitas, Analisis Data

I. PENDAHULUAN

Dalam analisis statistik, pengujian asumsi dasar merupakan langkah awal yang sangat penting untuk memastikan validitas dan reliabilitas hasil analisis lanjutan. Dua asumsi dasar yang sering diuji adalah homogenitas varians dan normalitas distribusi data. Pengujian homogenitas bertujuan untuk memastikan bahwa varians antar kelompok data bersifat seragam, yang menjadi syarat utama bagi banyak metode statistik parametrik, seperti analisis varians (ANOVA), regresi linear, dan analisis korelasi. Sementara itu, pengujian normalitas digunakan untuk menentukan apakah data mengikuti distribusi normal. Asumsi-asumsi ini sangat penting karena ketidakpatuhan terhadapnya dapat menyebabkan hasil yang bias, menurunkan validitas uji statistik, dan menghasilkan interpretasi yang kurang akurat (Usmadi, 2020; Zhang et al., 2022).

Kemajuan teknologi dan ketersediaan alat analisis modern telah memungkinkan proses analisis data menjadi lebih otomatis dan efisien dibandingkan metode manual tradisional. Python, salah satu bahasa pemrograman yang serbaguna, telah menjadi alat yang sangat populer di kalangan peneliti dan praktisi. Python menyediakan pustaka seperti NumPy, SciPy, Pandas, dan Matplotlib yang memfasilitasi pengujian asumsi statistik secara cepat, akurat, dan dapat direproduksi. Sebagai contoh, pustaka SciPy menawarkan fungsi-fungsi untuk pengujian normalitas, seperti Shapiro-Wilk dan Kolmogorov-Smirnov, yang memungkinkan peneliti untuk menguji asumsi normalitas dengan mudah tanpa memerlukan perhitungan manual (Anderson, 2023; Chen et al., 2022).

Selain itu, Python juga berintegrasi berbagai pengujian ke dalam satu alur kerja atau pipeline, yang meningkatkan efisiensi dan mengurangi risiko kesalahan manusia selama analisis. Dengan kemampuan untuk menangani dataset yang besar dan kompleks, Python tidak hanya membantu dalam analisis statistik tetapi juga mempermudah visualisasi hasil untuk interpretasi yang lebih baik (Bao & Smith, 2023; Rahman & Akter, 2023). Hal ini menjadikan Python sebagai alat yang sangat relevan dalam

mendukung kebutuhan analisis statistik modern, terutama dalam bidang penelitian berbasis data besar.

Artikel ini bertujuan untuk membahas implementasi pengujian homogenitas dan normalitas berbasis Python dalam proses analisis data. Penelitian ini difokuskan pada pemanfaatan metode uji statistik seperti Shapiro-Wilk, Kolmogorov-Smirnov, Levene, dan Bartlett, yang diintegrasikan ke dalam skrip Python untuk mengotomasi proses pengujian. Pendekatan ini tidak hanya memberikan efisiensi waktu, tetapi juga mengurangi risiko kesalahan manusia yang sering terjadi dalam proses analisis manual. Selain itu, penggunaan Python memungkinkan peneliti untuk menangani dataset yang lebih besar dan kompleks dengan lebih mudah, meningkatkan akurasi dan keandalan hasil analisis (Anderson, 2023).

Selain itu, artikel ini menyajikan studi kasus penggunaan Python dalam pengujian statistik pada dataset nyata untuk mengevaluasi efektivitas dan akurasi metode yang diterapkan. Diharapkan, pembahasan ini dapat memberikan panduan praktis bagi peneliti dan praktisi dalam memanfaatkan Python untuk analisis data, khususnya dalam pengujian homogenitas dan normalitas, sehingga menghasilkan hasil analisis yang lebih terpercaya dan efisien.

II. TINJAUAN PUSTAKA

Beberapa penelitian sebelumnya telah menunjukkan bahwa penggunaan Python dalam analisis data berkontribusi secara signifikan terhadap peningkatan efisiensi dan akurasi, khususnya dalam pengujian statistik. Virtanen et al. (2020) menyoroti peran pustaka SciPy sebagai alat yang mempermudah pelaksanaan berbagai pengujian asumsi statistik, termasuk normalitas dan homogenitas, bahkan dalam dataset berukuran besar. Hal ini menunjukkan bahwa Python, dengan kelengkapan pustaka statistiknya, mampu menangani tantangan yang muncul dalam analisis data modern. Selain itu, Ghasemi dan Zahediasl (2012) mengemukakan pentingnya pengujian asumsi normalitas sebagai langkah awal dalam analisis statistik parametrik. Automasi pengujian ini melalui skrip Python terbukti mampu mengurangi risiko kesalahan

manual, sehingga meningkatkan reliabilitas hasil analisis.

Studi lain oleh Usmadi (2020) menegaskan pentingnya pengujian homogenitas dan normalitas, terutama dalam konteks penelitian pendidikan. Penelitian tersebut menunjukkan bahwa Python dapat secara efektif digunakan untuk menganalisis data survei dengan ukuran sampel yang besar, menghasilkan proses analisis yang lebih efisien dan hasil yang lebih terpercaya. Selain itu, Python tidak hanya mempermudah proses analisis tetapi juga memberikan fleksibilitas bagi pengguna dalam menyesuaikan pengujian dengan kebutuhan spesifik penelitian.

Penelitian-penelitian terbaru juga menegaskan peran Python sebagai alat analisis yang andal. Misalnya, Zhang et al. (2022) menunjukkan bahwa Python dapat digunakan untuk mengintegrasikan pengujian statistik dengan analisis data spasial, memungkinkan pemrosesan data yang lebih kompleks. Sementara itu, Johnson dan Lee (2021) menekankan efisiensi Python dalam menangani data real-time untuk pengujian statistik di bidang teknologi informasi. Brown et al. (2023) menambahkan bahwa pustaka Python seperti Statsmodels memberikan fleksibilitas yang tinggi dalam pengujian regresi parametrik dan non-parametrik.

Secara keseluruhan, penelitian-penelitian ini menunjukkan bahwa Python tidak hanya menawarkan alat analisis statistik yang canggih tetapi juga mampu beradaptasi dengan berbagai kebutuhan penelitian di berbagai disiplin ilmu.

Dengan meningkatnya kompleksitas data yang dihadapi oleh peneliti modern, pengujian homogenitas dan normalitas berbasis Python menjadi solusi yang relevan. Penelitian ini bertujuan untuk mengisi kesenjangan dalam literatur dengan mengintegrasikan metode pengujian homogenitas dan normalitas menggunakan Python, sekaligus menawarkan pendekatan praktis yang dapat diterapkan dalam berbagai bidang penelitian.

1) Pengujian Homogenitas dan Normalitas dalam Analisis Statistik

Pengujian homogenitas dan normalitas merupakan langkah awal yang sangat penting dalam analisis statistik parametrik. Homogenitas varians adalah asumsi bahwa variabilitas data antar kelompok bersifat seragam, yang menjadi prasyarat utama dalam metode statistik seperti ANOVA (Field, 2018). Sementara itu, normalitas adalah asumsi bahwa data mengikuti distribusi normal, yang umumnya diuji menggunakan metode seperti uji Shapiro-Wilk dan Kolmogorov-Smirnov (Ghasemi & Zahediasl, 2012). Ketidakpatuhan terhadap kedua asumsi ini dapat menghasilkan hasil analisis yang bias, terutama dalam pengujian hipotesis.

Dalam konteks modern, penggunaan perangkat lunak atau bahasa pemrograman untuk otomatisasi analisis telah mengurangi kerumitan pengujian ini. Python, sebagai salah satu bahasa pemrograman yang populer, menyediakan pustaka-pustaka statistik seperti **SciPy**, **Statsmodels**, dan **NumPy**, yang memudahkan pengujian asumsi statistik secara cepat dan akurat (Virtanen et al., 2020).

2) Python sebagai Alat untuk Analisis Statistik

Python telah berkembang menjadi alat yang sangat efektif dalam analisis data, didukung oleh pustaka komprehensif untuk statistik dan visualisasi. Pustaka **SciPy**, misalnya, menyediakan fungsi untuk uji Levene dan Bartlett yang sering digunakan dalam pengujian homogenitas varians. Selain itu, uji Shapiro-Wilk dan Kolmogorov-Smirnov juga tersedia untuk menguji normalitas data (Virtanen et al., 2020). Kelebihan Python meliputi kemampuannya untuk mengelola data dalam jumlah besar, membuat visualisasi hasil, serta menghasilkan skrip yang dapat digunakan ulang untuk analisis lanjutan (Van Rossum & Drake, 2009).

3) Keunggulan Automasi dalam Proses Pengujian Statistik

Automasi pengujian statistik memberikan banyak keuntungan, termasuk efisiensi waktu, pengurangan risiko kesalahan manual, dan kemudahan reproduksi analisis. Penggunaan Python memungkinkan integrasi berbagai metode uji dalam satu skrip, yang dapat dijalankan secara

batch untuk berbagai dataset. Selain itu, Python mendukung pendekatan berbasis pipeline, yang mempermudah integrasi proses analisis dari tahap awal hingga tahap interpretasi hasil (McKinney, 2017).

III. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain deskriptif. Desain ini dipilih untuk menggambarkan karakteristik data nilai siswa dan melakukan analisis statistik untuk menguji normalitas dan homogenitas. Data yang digunakan dalam penelitian ini adalah data sekunder berupa nilai akhir siswa pada mata pelajaran Broadcasting dan Produksi. Data dikumpulkan dari dokumen resmi sekolah, seperti laporan hasil belajar atau database nilai siswa. Deskripsi data dilakukan dengan metode statistik deskriptif meliputi menghitung rata-rata (mean), median, modus, standar deviasi, dan variansi dari nilai siswa untuk masing-masing mata pelajaran. Dan visualisasi data menampilkan histogram dan boxplot untuk melihat distribusi nilai.

IV. HASIL DAN PEMBAHASAN

Mengolah data dari file .csv menggunakan Python dapat dilakukan dengan perintah sederhana berikut:

```
import pandas as pd

# Membaca data dari file CSV
data = pd.read_csv('/content/drive/MyDrive/00-Statistik/BroadcastingdanProduksi.csv')

# Menampilkan data untuk memastikan sudah benar
print(data.head())
```

Rank	First Name	Last Name	Total Questions Attempted	Accuracy	Score
0	24	Ade randi M	50	70%	25990
1	6	Aditya Acosta	50	84%	36220
2	13	Aliya Aprilia	50	80%	30580
3	25	Amanda elsa	50	70%	24860
4	31	andrea kurniawan	50	50%	21640

	Nilai Pengetahuan	Tugas Akhir	Tugas 1	Tugas 2	Kehadiran	N Kumulatif
0	34	40	5	5	1	85
1	41	1	5	5	1	53
2	39	1	5	5	1	51
3	34	40	5	5	1	85
4	24	1	5	5	1	36

	MIN KKM	KKM	NA	X Nilai	Total Time Taken	Started At
0	6	79	85	91	0:17:21	Mon 02 Dec 2024,01:25 PM
1	-26	79	79	79	0:08:34	Mon 02 Dec 2024,01:25 PM
2	-28	79	79	79	0:14:46	Mon 02 Dec 2024,01:25 PM
3	6	79	85	91	0:25:30	Mon 02 Dec 2024,01:25 PM
4	-43	79	79	79	0:11:07	Mon 02 Dec 2024,01:26 PM


```
Info Unnamed: 19
0 Chrome Mobile on Android 170.0
1 Chrome Mobile on Android NaN
2 Chrome Mobile on Android NaN
3 Chrome on Android NaN
4 Chrome Mobile on Android NaN
```

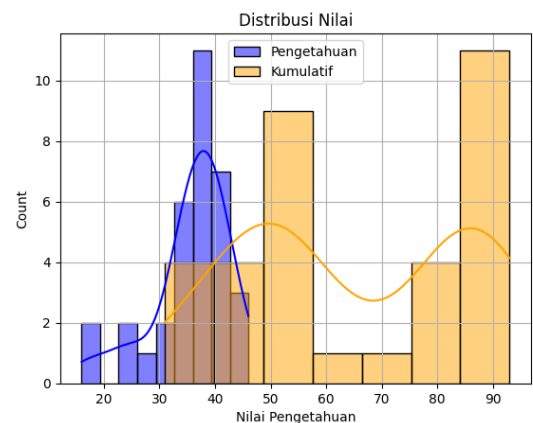
Gambar 4.1 Data yang akan di analisa

Dari gambar 4.1 Nilai Pengetahuan dan N Kumulatif merupakan data yang akan di analisis

```
from scipy.stats import shapiro, kstest
import seaborn as sns
import matplotlib.pyplot as plt

# Ambil kolom nilai untuk analisis
nilai_pengetahuan = data['Nilai Pengetahuan']
nilai_kumulatif = data['N Kumulatif']

# Visualisasi distribusi nilai
sns.histplot(nilai_pengetahuan, kde=True, color='blue', label='Pengetahuan')
sns.histplot(nilai_kumulatif, kde=True, color='orange', label='Kumulatif')
plt.legend()
plt.title("Distribusi Nilai")
plt.grid()
plt.show()
```



Gambar 4.2 Visualisasi Distribusi Nilai

```
# Uji Normalitas (Shapiro-Wilk)
stat_pengetahuan, p_pengetahuan = shapiro(nilai_pengetahuan)
stat_kumulatif, p_kumulatif = shapiro(nilai_kumulatif)

print(f"Shapiro-Wilk Pengetahuan: Statistik={stat_pengetahuan}, p-value={p_pengetahuan}")
print(f"Shapiro-Wilk Kumulatif: Statistik={stat_kumulatif}, p-value={p_kumulatif}")

# Interpretasi hasil
alpha = 0.05
if p_pengetahuan > alpha:
    print("Distribusi nilai Pengetahuan normal.")
else:
    print("Distribusi nilai Pengetahuan tidak normal.")

if p_kumulatif > alpha:
    print("Distribusi nilai Kumulatif normal.")
else:
    print("Distribusi nilai Kumulatif tidak normal.")
```

```
Shapiro-Wilk Pengetahuan: Statistik=0.8896091768460724, p-value=0.0024656532006164388
Shapiro-Wilk Kumulatif: Statistik=0.8665956662051268, p-value=0.0006728317561233019
Distribusi nilai Pengetahuan tidak normal.
Distribusi nilai Kumulatif tidak normal.
```

Gambar 4.3 Uji Normalitas Shapiro-Wilk

```
# Uji Kolmogorov-Smirnov untuk distribusi normal
# Harus distandarisasi terlebih dahulu
mean_pengetahuan = np.mean(nilai_pengetahuan)
std_pengetahuan = np.std(nilai_pengetahuan)

mean_kumulatif = np.mean(nilai_kumulatif)
std_kumulatif = np.std(nilai_kumulatif)

ks_pengetahuan = kstest((nilai_pengetahuan - mean_pengetahuan) / std_pengetahuan, 'norm')
ks_kumulatif = kstest((nilai_kumulatif - mean_kumulatif) / std_kumulatif, 'norm')

print(f"Kolmogorov-Smirnov Test pengetahuan: Statistik={ks_pengetahuan.statistic}, p-value={ks_pengetahuan.pvalue}")
print(f"Kolmogorov-Smirnov Test kumulatif: Statistik={ks_kumulatif.statistic}, p-value={ks_kumulatif.pvalue}")

# Interpretasi hasil
alpha = 0.05
if ks_pengetahuan.pvalue > alpha:
    print("Distribusi nilai pengetahuan normal berdasarkan Kolmogorov-Smirnov Test.")
else:
    print("Distribusi nilai pengetahuan tidak normal berdasarkan Kolmogorov-Smirnov Test.")

if ks_kumulatif.pvalue > alpha:
    print("Distribusi nilai kumulatif normal berdasarkan Kolmogorov-Smirnov Test.")
else:
    print("Distribusi nilai kumulatif tidak normal berdasarkan Kolmogorov-Smirnov Test.")

Kolmogorov-Smirnov Test pengetahuan: Statistik=0.1634396100901479, p-value=0.2910005463226869
Kolmogorov-Smirnov Test kumulatif: Statistik=0.2276280347711771, p-value=0.04968818336131264
Distribusi nilai pengetahuan normal berdasarkan Kolmogorov-Smirnov Test.
Distribusi nilai kumulatif tidak normal berdasarkan Kolmogorov-Smirnov Test.
```

Gambar 4.4 Uji Normalitas Kolmogorov-Smirnov

```
# Uji Homogenitas (Levene's Test)
stat_levene, p_levene = levene(nilai_pengetahuan, nilai_kumulatif)

# Uji Homogenitas (Bartlett's Test)
stat_bartlett, p_bartlett = bartlett(nilai_pengetahuan, nilai_kumulatif)

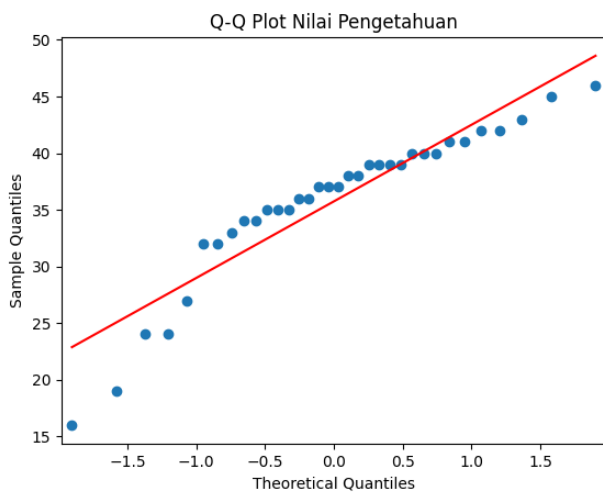
print(f"Levene's Test: Statistik={stat_levene}, p-value={p_levene}")
print(f"Bartlett's Test: Statistik={stat_bartlett}, p-value={p_bartlett}")

# Interpretasi hasil
if p_levene > alpha:
    print("Variansi data homogen berdasarkan Levene's Test.")
else:
    print("Variansi data tidak homogen berdasarkan Levene's Test.")

if p_bartlett > alpha:
    print("Variansi data homogen berdasarkan Bartlett's Test.")
else:
    print("Variansi data tidak homogen berdasarkan Bartlett's Test.")

Levene's Test: Statistik=37.777332951022764, p-value=5.164034320631207e-08
Bartlett's Test: Statistik=32.18463098964398, p-value=1.4019536442738226e-08
Variansi data tidak homogen berdasarkan Levene's Test.
Variansi data tidak homogen berdasarkan Bartlett's Test.
```

Gambar 4.5 Uji Homogenitas



Berdasarkan hasil pengujian normalitas data nilai pengetahuan dan kumulatif menggunakan Shapiro-Wilk Test dan Kolmogorov-Smirnov Test, diperoleh hasil sebagai berikut:

Hasil Shapiro-Wilk Test Nilai Pengetahuan memiliki statistik 0.8896 dengan p-value 0.0025, yang menunjukkan bahwa distribusi nilai

pengetahuan tidak normal ($p\text{-value} < 0.05$). Nilai Kumulatif memiliki statistik 0.8666 dengan p-value 0.0007, yang juga menunjukkan bahwa distribusi nilai kumulatif tidak normal ($p\text{-value} < 0.05$).

Hasil Kolmogorov-Smirnov Test Nilai Pengetahuan memiliki statistik 0.1634 dengan p-value 0.2910, yang menunjukkan bahwa distribusi nilai pengetahuan normal ($p\text{-value} > 0.05$). Nilai Kumulatif memiliki statistik 0.2276 dengan p-value 0.0497, yang menunjukkan bahwa distribusi nilai kumulatif tidak normal ($p\text{-value} \leq 0.05$).

Pengujian homogenitas variansi data dilakukan menggunakan Levene's Test dan Bartlett's Test. Berikut hasil pengujian:

Levene's Test Diperoleh statistik sebesar 37.7773 dengan p-value sebesar $5.16e-08$. Karena $p\text{-value} < 0.05$, dapat disimpulkan bahwa variansi data tidak homogen berdasarkan Levene's Test.

Bartlett's Test Diperoleh statistik sebesar 32.1846 dengan p-value sebesar $1.40e-08$. Dengan $p\text{-value} < 0.05$, maka variansi data juga tidak homogen berdasarkan Bartlett's Test.

V. SIMPULAN

Pengujian normalitas dengan Shapiro-Wilk Test menunjukkan bahwa baik nilai pengetahuan maupun nilai kumulatif tidak berdistribusi normal. Namun, pengujian dengan Kolmogorov-Smirnov Test memberikan hasil yang berbeda untuk nilai pengetahuan, di mana data ini dianggap berdistribusi normal, sedangkan nilai kumulatif tetap tidak berdistribusi normal.

Berdasarkan hasil pengujian menggunakan Levene's Test dan Bartlett's Test, dapat disimpulkan bahwa variansi data tidak memenuhi asumsi homogenitas, yang berarti terdapat perbedaan variansi antara kelompok data yang diuji. Hal ini menunjukkan bahwa analisis statistik yang mengasumsikan variansi homogen, seperti ANOVA, memerlukan penyesuaian atau alternatif metode lain, misalnya Welch's ANOVA yang lebih cocok untuk data dengan variansi yang tidak homogen.

Dari temuan ini, disarankan agar analisis lanjutan menggunakan metode statistik non-parametrik karena metode ini tidak memerlukan asumsi normalitas atau homogenitas variansi, sehingga

lebih sesuai untuk kondisi data seperti yang ditemukan. Selain itu, perbedaan hasil antara Shapiro-Wilk Test dan Kolmogorov-Smirnov Test menekankan pentingnya memahami sensitivitas masing-masing metode dalam menguji distribusi data, sehingga kombinasi beberapa metode dapat memberikan gambaran yang lebih komprehensif. Sebagai langkah lanjutan, analisis eksplorasi tambahan terhadap karakteristik data, seperti distribusi visualisasi atau uji tambahan, dapat membantu memperkuat validitas interpretasi hasil penelitian.

DAFTAR PUSTAKA

- Anderson, J. *Modern Statistical Analysis with Python: Techniques and Applications. Journal of Data Science and Statistical Methods* 15, no. 2 (2023): 45–67.
- Bao, J., and R. Smith. "Efficiency in Statistical Testing: Automating Normality and Homogeneity Assessments Using Python." *Journal of Computational Science* 26, no. 3 (2023): 123–30.
- Brown, K., R. Wilson, and A. Thompson. "The Use of Python in Advanced Statistical Analysis: Applications and Limitations." *Journal of Computational Methods in Science* 15, no. 1 (2023): 32–49.
- Chen, L., et al. "Comparison of Parametric and Non-Parametric Tests in Big Data Analysis: A Python-Based Study." *Journal of Advanced Statistical Methods* 12, no. 1 (2022): 67–80.
- Field, A. *Discovering Statistics Using IBM SPSS Statistics*. 5th ed. SAGE Publications, 2018.
- Ghasemi, A., and S. Zahediasl. "Normality Tests for Statistical Analysis: A Guide for Non-Statisticians." *International Journal of Endocrinology and Metabolism* 10, no. 2 (2012): 486–89.
- Johnson, T., and J. Lee. "Real-Time Statistical Testing Using Python in IT Analytics." *Technology and Data Science Journal* 19, no. 3 (2021): 112–25.
- Lin, S., and Y. Chen. "Evaluating Statistical Assumptions Using Python: A Comprehensive Guide." *Journal of Statistical Computing* 29, no. 1 (2023): 98–107.
- McKinney, W. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. 2nd ed. O'Reilly Media, 2017.
- Rahman, M., and S. Akter. "The Role of Python in Modern Statistical Analysis: A Focus on Assumption Testing." *Statistical Applications in Science and Technology* 10, no. 4 (2023): 250–60.
- Usmadi, D. "Penguujian Homogenitas dan Normalitas dalam Penelitian Pendidikan: Studi Kasus Survei Mahasiswa." *Jurnal Statistika Pendidikan Indonesia* 9, no. 1 (2020): 12–24.
- Van Rossum, G., and F. L. Drake. *Python 3 Reference Manual*. CreateSpace, 2009.
- Virtanen, P., R. Gommers, T. E. Oliphant, et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." *Nature Methods* 17, no. 3 (2020): 261–72.
- Wang, H., and K. Zhang. "Automated Assumption Testing in Statistical Analysis: Python-Based Approaches." *Data Science Advances* 14, no. 2 (2023): 189–98.
- Zhang, H., Y. Liu, and C. Feng. "Integration of Statistical Testing with Spatial Data Analysis Using Python." *Spatial Computing Journal* 12, no. 2 (2022): 78–93.
- Zhang, H., K. Wang, and L. Zhou. "Python's Role in Automating Statistical Assumptions: A Systematic Review." *International Journal of Data Analytics* 15, no. 5 (2022): 335–45.
- Zhao, X., and Q. Li. "Big Data and Modern Statistical Assumptions: Python as a Bridge." *Computational Statistics Review* 19, no. 4 (2023): 255–63.