

Penerapan Anova One Way Untuk Menganalisis Pengaruh Pendidikan Terhadap Pendapatan Individu Menggunakan Python Serta Visualisasi Data Dengan Looker Studio

Woro Isti Rahayu^{1)*}, Fatia Amalia Maresti²⁾, Aulia Rahmadiva Wardana³⁾, Firna Ulfiani Ramadhani⁴⁾

^{1,2,3,4)}Universitas Logistik dan Bisnis Internasional, Fakultas Logistik, Teknologi dan Bisnis, Program Studi Sains Data

woroiisti@ulbi.ac.id¹⁾, fatiaamalia@ulbi.ac.id²⁾, aularahmadiva4949@gmail.com³⁾, firnaramadhani2121@gmail.com⁴⁾

Received: 20 January 2025

Accepted: 3 February 2025

Published: 6 February 2025



*woroiisti@ulbi.ac.id

Kata Kunci: Pendidikan, Pendapatan, Anova One Way, Looker Studio

DSI: Jurnal Data Science Indonesia is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak : Penelitian ini bertujuan untuk mengetahui seberapa besar Pengaruh Tingkat Pendidikan terhadap Pendapatan Individu. Selain itu juga meneliti pendapatan individu jika dilihat dari usia, jenis kelamin, kepemilikan sertifikasi/keahlian, jenis pekerjaan, dan lama bekerja. Metode analisis yang dipakai di dalam penelitian ini adalah metode ANOVA one way karena dianggap paling akurat dalam menyelesaikan masalah ini. Sumber data yang digunakan adalah data yang dikumpulkan dari survey online menggunakan Google Form. Hasil penelitian mendapatkan bahwa beberapa tingkat pendidikan berpengaruh signifikan terhadap tingkat pendapatan individu, dimana semakin tinggi tingkat pendidikan maka tingkat pendapatan juga akan meningkat. Faktor pendukung seperti kepemilikan sertifikasi/keahlian juga berpengaruh terhadap beberapa tingkat pendidikan. Selain itu, Pada visualisasi data menggunakan looker studio pengaruh lama bekerja dan usia tergantung kepada jenis pekerjaan, dalam penelitian ini, pada wiraswasta dan pedagang lama bekerja dan usia tidak mempengaruhi jumlah pendapatan, tetapi pada petani faktor lama bekerja mempengaruhi jumlah pendapatan

PENDAHULUAN

Faktor penting dalam investasi sumberdaya manusia adalah pendidikan. Pendidikan merupakan faktor penting dalam investasi sumber daya manusia sehingga pendidikan juga menjadi pengaruh dalam tingkat pendapatan[1]. Dengan pendidikan individu dapat meningkatkan taraf hidup yang mana hal tersebut pastinya berhubungan dengan pendapatan bersih. Pendapatan masyarakat menjadi hal yang juga diperhatikan pemerintah, saat ini indonesia menjadi negara menengah bawah (*Lower Middle-Income Country*) jika dihitung dari pendapatan per kapita nya[2]. Oleh sebab itu, semestinya pemerintah memberikan perhatian lebih terhadap pendidikan di Indonesia dan masyarakat juga meningkatkan kesadarannya akan pendidikan karena saat ini pendidikan yang menjadi kunci utama dalam penentuan penghasilan.

Tingkat pendidikan mempunyai pengaruh yang signifikan terhadap kualitas tenaga kerja. Oleh karena itu, Tingkat pendidikan yang lebih tinggi cenderung menentukan kualitas sumber daya manusia di dunia kerja[3]. Ada 3 tingkatan pendidikan di Indonesia yaitu Sekolah Dasar (SD) atau sederajat yang ditempuh selama 6 tahun, Sekolah Menengah Pertama (SMP) atau sederajat yang selama 3 tahun, dan Sekolah Menengah Atas atau sederajat (SMA) atau sederajat yang ditempuh selama 3 tahun. Tetapi, pada saat ini banyak industri yang menginginkan pekerja dengan minimal tamatan SMA, dimana hal itu membuat terjadinya perbedaan pendapatan di antara individu di Indonesia[4].

Penelitian ini bertujuan untuk menganalisis apakah terdapat perbedaan yang signifikan pada rata-rata pendapatan berdasarkan tingkat pendidikan. Penelitian ini juga mengkaji pengaruh faktor-faktor penunjang lainnya, seperti sertifikasi/keahlian, usia, lama bekerja, jenis kelamin, dan jenis pekerjaan terhadap jumlah pendapatan pada setiap tingkat pendidikan. Dalam penelitian ini, metode *One-Way ANOVA* digunakan untuk

menguji perbedaan rata-rata pendapatan pada masing-masing tingkat pendidikan dan menguji dan pengaruh faktor penunjang keahlian/sertifikasi pada setiap tingkat pendidikan yang ada.

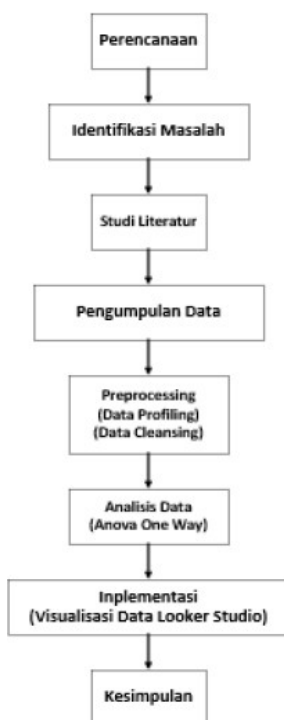
Selain itu, hasil visualisasi dibuat untuk melihat pengaruh factor penunjang usia, lama bekerja, jenis kelamin, dan jenis pekerjaan terhadap jumlah pendapatan pada setiap tingkat pendidikan yang ada melalui *dashboard* pada *Looker Studio*. Oleh karena itu penelitian ini berjudul "Penerapan *Anova One Way* Untuk Menganalisis Pengaruh Tingkat Pendidikan Terhadap Pendapatan Individu Menggunakan Python Serta Visualisasi Data Dengan Looker Studio".

TINJAUAN LITERATUR

Penelitian sebelumnya dilakukan oleh Yeni Setiani,dkk dengan judul "Visualisasi Data Malnutrisi Anak Di Asia Menggunakan Looker Studio Serta Analisis Data Dengan Metode ANOVA" dalam penelitian ini didapatkan adanya pengaruh yang signifikan antara klasifikasi pendapatan negara dan persentase malnutrisi anak yang terjadi di negara tersebut, bukan hanya terjadi secara kebetulan serta memberikan wawasan penting tentang masalah malnutrisi anak di Asia dalam bentuk visual[5]. Selain itu penelitian lain juga dilakukan oleh Chelsey Hill, dkk dengan judul "*Comparing programming languages for data analytics: Accuracy of estimation in Python and R*", hasilnya menunjukkan bahwa paket stats di R dan pustaka statsmodels di Python dapat diandalkan untuk statistik ringkasan univariat. Sebaliknya, pustaka scikit-learn Python menghasilkan hasil yang paling akurat dan direkomendasikan untuk ANOVA[6]. Dengan fokus pada pendapatan dan pendidikan, penelitian kali ini menggunakan metode ANOVA ONE WAY untuk menganalisis pengaruh pendidikan terhadap pendapatan dengan menggunakan *library python* dan visualisasi data menggunakan *Looker Studio*.

METODE PENELITIAN

Penelitian ini bertujuan untuk menganalisis pengaruh tingkat pendidikan terhadap pendapatan individu dan faktor penunjang sertifikasi/keahlian terhadap pendapatan disetiap tingkat pendidikan menggunakan *Anova One Way*. Serta melihat pengaruh jenis kelamin, usia, dan lama bekerja menggunakan *dashboard Looker Studio*. Adapun alur penelitiannya adalah sebagai berikut.



Gambar 1 Diagram Alur Penelitian

1. Perencanaan

Menentukan akar permasalahan yang relevan saat ini yaitu "Pengaruh Pendidikan Terhadap Pendapatan Individu". Melakukan kajian pustaka tentang penelitian sebelumnya mengenai "Pengaruh Pendidikan Terhadap Pendapatan Individu". Menyusun hipotesis permasalahan yang diangkat yaitu "Pengaruh Pendidikan Terhadap Pendapatan Individu". Menentukan metode penelitian yang relevan dengan permasalahan yang diangkat. Memilih metode survei sebagai desain penelitian, dengan populasi pekerja berusia 20-50 tahun dan menggunakan kuesioner sebagai instrumen. Mengumpulkan data dari sampel yang representatif dan menganalisisnya dengan teknik statistik menggunakan Excel dan Bahasa Pemrograman Python

2. Identifikasi Masalah

Dari akar permasalahan yang didapatkan, masalah yang dapat diidentifikasi yaitu : Bagaimana pengaruh tingkat pendidikan terhadap pendapatan individu menggunakan *Anova One Way* dalam bahasa pemrograman python dan pada hasil visualisasi dengan Looker Studio?; Bagaimana pengaruh faktor penunjang seperti keahlian/sertifikasi, pada setiap tingkat pendidikan terhadap pendapatan individu menggunakan *Anova One Way* dalam bahasa pemrograman python ?; Bagaimana pengaruh faktor penunjang, seperti jenis kelamin, jenis pekerjaan, usia, dan lama bekerja terhadap pendapatan individu pada hasil visualisasi data menggunakan Looker Studio?

3. Studi Literatur

Melakukan kajian pustaka tentang penelitian sebelumnya mengenai "Pengaruh Pendidikan Terhadap Pendapatan Individu" .

4. Pengumpulan Data

Pada tahap ini data dikumpulkan diambil dari *Survey Online* dengan menyebarkan form yang telah dibuat untuk mendapatkan responden, untuk menganalisis pengaruh pendidikan terhadap pendapatan individu di indonesia. Melalui *survey* ini , didapatkan 72 responden.

5. Preprocessing

Tahapan ini merupakan tahapan yang dilakukan untuk mempersiapkan data agar data dapat dianalisis[7]. Adapun hal yang dilakukan kepada data yang telah dikumpulkan tadi yaitu :

a. Data Profiling sangat penting untuk memilih alat dan teknik yang akan gunakan untuk memproses data sesuai dengan karakteristiknya, dan memungkinkan kita untuk memahami data dengan lebih baik Pada tahap ini ada beberapa hal yang dilakukan yaitu : Pemeriksaan Struktur Data. Memeriksa struktur dataset yang didapat dari form, seperti jumlah baris dan kolom, jenis data dari setiap kolom (misalnya, numerik, kategorikal, teks), serta hubungan antar kolom; Deteksi Data Hilang. Mengidentifikasi kolom atau baris yang memiliki data hilang (missing values) untuk menentukan strategi penanganan data hilang, misalnya dengan imputasi atau penghapusan.

b. Data Cleaning

Data Cleaning adalah perbaikan data yang rusak, hilang, atau tidak akurat (kesalahan). Pada fase pembersihan data, data harus dianalisis untuk mengidentifikasi outlier[8]. Outlier adalah kasus atau data dengan karakteristik unik yang tampak sangat berbeda dari pengamatan lain dan muncul sebagai nilai ekstrem baik pada variabel tunggal maupun kombinasi variabel. Pada tahap ini ada beberapa hal yang dilakukan, yaitu : Mengubah tipe data; Penghapusan kolom yang tidak diperlukan; Menangani missing value; Penghapusan karakter yang tidak diperlukan.

6. Analisis Data (ANOVA ONE WAY)

Anova atau Analysis of variance (anova) adalah tergolong analisis komparatif dua variabel atau lebih dari dua rata-rata. Tujuannya untuk membandingkan lebih dari dua rata-rata. Gunanya untuk menguji kemampuan generalisasi artinya data sampel dianggap mewakili populasi[9].



Gambar 2 Diagram Alur ANOVA ONE WAY

- Perumusan Hipotesis: Hipotesis Nol (H_0): Tidak ada perbedaan yang signifikan antara rata-rata persentase pengaruh setiap tingkat pendidikan terhadap pendapatan individu. Hipotesis Alternatif (H_1): Terdapat perbedaan yang signifikan antara rata-rata persentase pengaruh setiap tingkat pendidikan terhadap pendapatan individu
- Uji Asumsi Normalitas
Uji normalitas adalah suatu teknik statistik untuk menentukan apakah data atau sampel populasi memiliki distribusi normal atau mendekati normal[10]. Selain menggunakan rumus uji Shapiro Wilk, pada uji normalitas kami juga menggunakan *library scipy.stats* sehingga jika menghasilkan p-value lebih besar dari tingkat signifikansi (0,05) maka dapat dinyatakan bahwa Hipotesis 0 diterima dan Hipotesis Alternatif ditolak dan sebaliknya jika hasil uji normalitas menghasilkan p-value lebih kecil dari tingkat signifikansi (0,05) maka dapat dinyatakan bahwa Hipotesis 0 ditolak dan Hipotesis Alternatif diterima. Pengujian asumsi normalitas juga dapat dilakukan visual menggunakan *probability plot* dengan *library matplotlib.pyplot*. Berikut gambar 4 merupakan library yang akan digunakan untuk menguji Asumsi Normalitas dan akan diolah pada lingkungan kerja *Google Collab*.

```

import scipy.stats as stats

import matplotlib.pyplot as plt
  
```

Gambar 3 Library Uji Asumsi Normalitas

- Uji Asumsi Varians
Selain menggunakan uji Levene untuk memastikan apakah homogenitas varians[11], pada Uji Asumsi Varians ini digunakan library pingouin dengan hasil p-value sebagai nilai uji levene. pada python Berikut library pingouin yang kami gunakan untuk menguji dataset.

```

import pingouin as pg
pg.homoscedasticity()
  
```

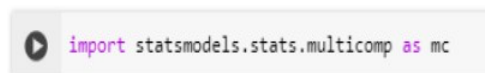
Gambar 3 Library Library Asumsi Varians

- ANOVA
Uji ANOVA digunakan untuk menentukan apakah perbedaan rata-rata yang signifikan ditemukan dalam lebih dari dua kelompok sampel. Ada beberapa library yang dapat digunakan diantaranya : statsmodels, scipy, pingouin. Dengan aturan jika p-value yang didapat lebih kecil dibandingkan

tingkat signifikansi (0,05) yang digunakan maka dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara tiap kategori, tetapi jika sebaliknya maka artinya tidak ada perbedaan yang signifikan antara kategorinya. Untuk menemukan p-value yang menunjukkan signifikansi statistik dari pengujian hipotesis nol bahwa tidak ada pengaruh dari variabel independen, lihat tabel anova. Nilai p-value yang lebih kecil menunjukkan bahwa pengaruh variabel independen lebih signifikan[12]

- Uji Post Hoc (Tukey)

Pengujian Turkey merupakan suatu pengujian perbandingan jamak untuk menentukan perbedaan antara tiga atau lebih kelompok obyek penelitian[13]. Selain menggunakan rumus uji Post Hoc (Tukey), berikut kami juga menggunakan *library statsmodels* pada python untuk menguji dataset.



Gambar 5 Library Uji Post Hoc (Tukey)

- Uji Homogeneous Subsets

Subset homogen merujuk pada kelompok rerata yang tidak berbeda secara signifikan satu sama lain berdasarkan uji post hoc. Subset ini dihasilkan ketika seorang peneliti melakukan beberapa perbandingan setelah hasil ANOVA yang signifikan untuk mengidentifikasi kelompok spesifik mana yang berbeda[14]. Pada uji ini kita dapat mengetahui kategori mana yang mempunyai perbedaan yang paling signifikan diantara kategori yang lain.

7. Implementasi (Visualisasi Data dengan Looker Studio)

Data visualisasi merupakan representasi grafis dari data yang bertujuan untuk membantu pemahaman, analisis dan komunikasi data supaya lebih efektif. Looker studio adalah alat Business Intelligence (BI) yang berfungsi untuk mengumpulkan, menganalisis, dan memvisualisasikan informasi[15]. Data visualisasi ditampilkan dalam bentuk dashboard menggunakan berbagai jenis grafik, diagram dan visual lainnya untuk menggambarkan pola, hubungan, tren dan informasi data penting dalam data.[16] Hasil dari visualisasi ini adalah grafik yang efektif dan menarik sesuai dengan data yang diberikan dan dapat dijalankan pada platform seluler. Pada tahap ini, dilakukan proses visualisasi data menggunakan looker studio yang bertujuan untuk menggambarkan data tingkat pendidikan terhadap pendapatan individu di Indonesia. Tahapan visualisasi data pada looker data studi yaitu : Mempersiapkan data yang akan dianalisis; Mengkoneksikan sumber data dan Looker Studio; Menyajikan grafik/diagram yang dibutuhkan atau yang ingin ditampilkan

8. Kesimpulan

Menafsirkan hasil dari analisis menggunakan metode *Anova One Way* dan hasil visualisasi menggunakan Looker Studio.

HASIL PENELITIAN

a. Pengumpulan Data

Pada tahap ini data dikumpulkan diambil dari *Survey Online* dengan menyebarkan form yang telah dibuat untuk mendapatkan responden, untuk menganalisis pengaruh pendidikan terhadap pendapatan individu di Indonesia. Melalui *survey* ini kami berhasil mengumpulkan 72 responden.

b. Preprocessing

Tahapan ini merupakan tahapan yang dilakukan untuk mempersiapkan data agar data dapat dianalisis. Adapun hal yang dilakukan kepada data yang telah dikumpulkan tadi yaitu :

1. Data Profiling

Pada tahap ini ada beberapa hal yang dilakukan yaitu :

- Pemeriksaan Struktur Data. Peneliti memeriksa struktur dataset yang didapat dari form, seperti jumlah baris dan kolom, jenis data dari setiap kolom (misalnya, numerik, kategorikal, teks), serta hubungan antar kolom.
- Deteksi Data Hilang. Peneliti mengidentifikasi kolom atau baris yang memiliki data

hilang (missing values) untuk menentukan strategi penanganan data hilang, misalnya dengan imputasi atau penghapusan.

2. Data Cleaning

Pada tahap ini ada beberapa hal yang dilakukan, yaitu : Mengubah tipe data; Penghapusan kolom yang tidak diperlukan; Menangani missing value; Penghapusan karakter yang tidak diperlukan

c. Analisis Data Anova One Way Pada Pengaruh Tingkat Pendidikan Terhadap Pendapatan Menggunakan Bahasa Pemrograman *Python*

Penggunaan bahasa pemrograman *python* sebagai otomatisasi penerapan anova one way agar pengerjaannya dapat dilakukan dengan lebih efisien. Berikut merupakan langkah-langkah penerapan anova *one way* menggunakan *python* untuk mengetahui pengaruh dan perbedaan signifikan antara rata-rata pendapatan disetiap tingkat pendidikan :

1. Uji Asumsi Normalitas

Sebelum melakukan uji asumsi normalitas terlebih dahulu mencari nilai residual dengan *function ols*, seperti dibawah ini

```
[ ] import statsmodels.api as sm
    from statsmodels.formula.api import ols
    model_pendidikan = ols('Pendapatan ~ C(Pendidikan)', data=data_bersih).fit()
    residual = model_pendidikan.resid
```

Gambar 6 Melakukan uji asumsi normalitas

Dari hasil output yang diberikan pada (gambar 6), nilai statistik uji Wilk sebesar 0.91 dan nilai p-value sebesar 0.00. Dengan menggunakan taraf signifikansi 5% disimpulkan berdasarkan nilai p-value residual sebesar 0.000 yang lebih kecil dari 0.05, menunjukkan bahwa hipotesis nol ditolak, sehingga dapat diartikan distribusi residual tidak normal. Dengan demikian, pada taraf signifikansi 5%, kita dapat menerima asumsi bahwa residual berdistribusi tidak normal. Karena *anova one way* merupakan metode statistik parametrik, mengharuskan asumsi normalitas harus terpenuhi. Untuk membuat data ini memenuhi asumsi normalitas yaitu berdistribusi normal, digunakan fungsi **np.log** seperti berikut:

```
import numpy as np
data_transformed = data_bersih.copy()
data_transformed['Pendapatan'] = np.log(data_bersih['Pendapatan'])
data_transformed
```

Gambar 7 Melakukan uji asumsi normalitas

Dapat dilihat pada (gambar 7) *p-value* yang dihasilkan yaitu $0,199 > 0,05$ (taraf signifikansi) menandakan bahwa tidak ada bukti yang cukup kuat untuk menolak hipotesis nol dan menyimpulkan bahwa data berdistribusi normal. Secara visual menggunakan grafik histogram, dapat terlihat juga bahwa residual berdistribusi normal. Karena asumsi normalitas telah terpenuhi, maka dapat dilanjutkan ke tahap selanjutnya yaitu uji varians.

2. Asumsi Varians

```
[58] import pingouin as pg
    homoscedasticity_result = pg.homoscedasticity(data_transformed, dv='Pendapatan', group='Pendidikan')
    homoscedasticity_result
```

	W	pval	equal_var
levene	1.254605	0.297777	True

Gambar 8 Melakukan uji asumsi normalitas

Pada asumsi varians, output yang dihasilkan pada (gambar 8) uji levene dapat disimpulkan bahwa pada taraf signifikansi 5%, residual memenuhi asumsi *variens* (h nol diterima). Hal ini dikarenakan residual memiliki *p-value* sebesar $0.297 > 0.05$. Karena asumsi *variens* telah dipenuhi maka dapat dilanjutkan ke tahap selanjutnya yaitu uji Anova *one way*.

3. Asumsi Anova

```
[59] import pingouin as pg

tabel_anov = pg.anova(data= data_transformed, dv = 'Pendapatan', between= 'Pendidikan', detailed=True)
tabel_anov
```

	Source	SS	DF	MS	F	p-unc	np2
0	Pendidikan	7.088567	3	2.362856	4.48127	0.006533	0.178196
1	Within	32.690965	62	0.527274	NaN	NaN	NaN

Gambar 9 Hasil Uji Asumsi Anova

Pada (gambar 9) output diatas, dihasilkan nilai $p\text{-unc}$ (0,006) < nilai taraf signifikansi (0,05) sehingga dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara beberapa tingkat pendidikan.

4. Uji Post Hoc (Tukey)

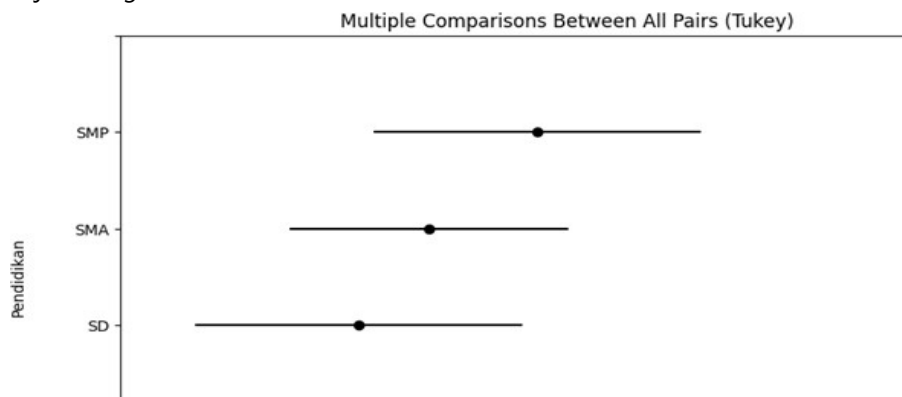
```
[63] import statsmodels.stats.multicomp as mc
comp = mc.MultiComparison(data_transformed['Pendapatan'],data_bersih['Pendidikan'])
post_hoc_res = comp.tukeyhsd()
post_hoc_res.summary()
```

Multiple Comparison of Means - Tukey HSD, FWER=0.05							
	group1	group2	meandiff	p-adj	lower	upper	reject
S1/D4/D3	SD		-0.8928	0.007	-1.5929	-0.1928	True
S1/D4/D3	SMA		-0.7408	0.0189	-1.3889	-0.0927	True
S1/D4/D3	SMP		-0.5091	0.2302	-1.2092	0.1909	False
SD		SMA	0.1521	0.9255	-0.496	0.8002	False
SD		SMP	0.3837	0.4754	-0.3163	1.0837	False
SMA		SMP	0.2316	0.7816	-0.4165	0.8797	False

Gambar 10 Hasil Uji Post Hoc (Tukey)

Pada (gambar 10) asumsi *Pos Hoc* diatas terpenuhi, maka maka multiple comparisons dilakukan dengan uji Tukey. Output dari pengujian diatas dapat disimpulkan bahwa: Ada perbedaan yang signifikan antara pendapatan per individu pada tingkat pendidikan S1/D3/D4 dan SD. Hal ini dapat dilihat dari nilai p-adj yang lebih kecil dari tingkat signifikansi yang digunakan 5% (0,007 0,05); Ada perbedaan yang signifikan antara pendapatan per individu pada tingkat pendidikan S1/D3/D4 dan SMA, dapat dilihat dari nilai p-adj yang dihasilkan lebih kecil dari tingkat signifikansi yang digunakan 5% (0,0189 0,05); Tidak ada perbedaan yang signifikan antara pendapatan per individu pada tingkat pendidikan S1/D3/D4 dan SMP, dapat dilihat dari nilai p-adj yang dihasilkan lebih besar dari tingkat signifikansi yang digunakan 5% (0,2302 0,05); Tidak ada perbedaan yang signifikan antara pendapatan per individu pada tingkat pendidikan SD dan SMA. Hal ini dapat dilihat dari nilai p-adj yang lebih besar dari tingkat signifikansi yang digunakan 5% (0,2955 0,05); Tidak ada perbedaan yang signifikan antara pendapatan per individu pada tingkat pendidikan SD dan SMP, dapat dilihat dari nilai p-adj yang dihasilkan lebih besar dari tingkat signifikansi yang digunakan 5% (0,4754 0,05); Tidak ada perbedaan yang signifikan antara pendapatan per individu pada tingkat pendidikan SMA dan SMP, dapat dilihat dari nilai p-adj yang dihasilkan lebih besar dari tingkat signifikansi yang digunakan 5% (0,7816 0,05).

5. Uji *Homogeneous Subsets*



Gambar 11 Hasil Uji *Homogeneous Subsets*

Kesimpulan dari (Gambar 11) homogeneous subset diatas adalah digunakan merujuk pada kelompok-kelompok data yang memenuhi asumsi homogenitas varians, sehingga dapat dianalisis menggunakan ANOVA one-way dengan keyakinan bahwa asumsi tersebut terpenuhi. Pada hasil uji diatas yaitu Multiple Comparisons Between All Pairs terlihat bahwa pendapatan tertinggi berada pada tingkat pendidikan S1/D4/D3 karena menempati urutan paling kanan.

Hasil akhir dari analisis Pendidikan terhadap pendapatan individu, didapat bahwa setiap tingkat pendidikan S1/D4/D3 dan SD sangat berpengaruh secara signifikan terhadap pendapatan individu.

- d. Hasil Dari Pengaruh Penunjang Lain Seperti Kepemilikan Keahlian/Sertifikasi Terhadap Pendapatan Individu Tingkat SD,SMP,SMA,S1/D4/D3

Dengan menggunakan metode yang sama, hasil pengujian tahap Anova One Way untuk pengaruh kepemilikan sertifikasi/keahlian pada tingkat pendidikan SD,SMP,SMA,S1/D4/D3 dapat disimpulkan bahwa uji *Multiple Comparisons Between All Pairs* didapatkan bahwa pendapatan pada tingkat pendidikan SD,SMP,SMA,S1/D4/D3 pendapatan tertinggi berada pada individu yang memiliki keahlian/sertifikasi dan ditandai dengan garis paling kanan. Tetapi tingkat pendidikan S1/D4/D3 yang memiliki pendapatan tertinggi diantara tingkat pendidikan lainnya.

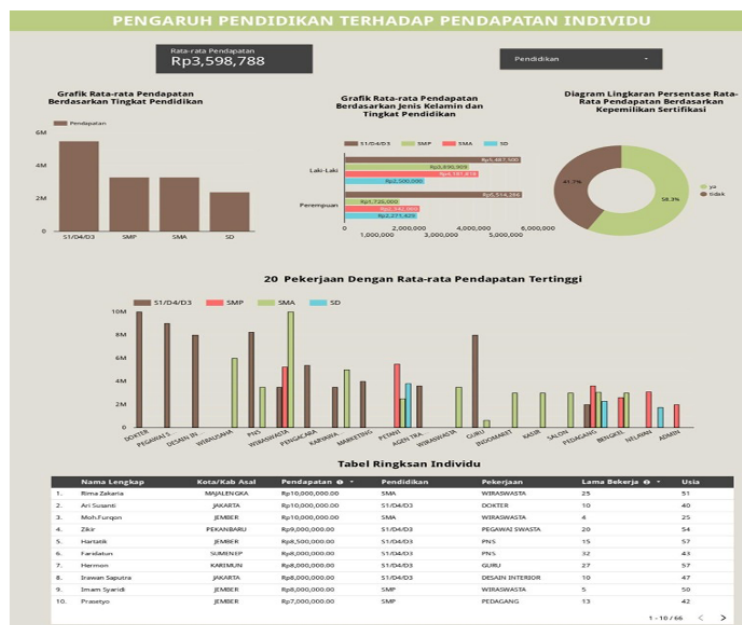
- e. Implementasi Menggunakan *Looker Studio*

Pada tahap pembuatan dashboard interaktif menggunakan Looker Studio, dimulai dengan pengunduhan data yang telah diproses di *Google Colab*, lalu menghubungkan sumber data ke *Looker Studio* menggunakan *Google Sheets*. Setelah itu, berbagai grafik dan diagram dibuat, seperti *Column Chart*, *Donut Chart*, dan *Stacked Bar Chart*, dengan memanfaatkan fitur visualisasi untuk menggambarkan hubungan antara pendidikan, pendapatan, keahlian, jenis kelamin, pekerjaan, dan usia. Setiap grafik disesuaikan dengan format yang relevan, seperti pengaturan rata-rata pendapatan dan format mata uang IDR.

Setelah grafik dan tabel disusun, posisi visualisasi diatur agar tampilan dashboard lebih terstruktur. Tema dan warna dashboard disesuaikan agar tampak serasi, dan judul ditambahkan untuk memberikan informasi yang jelas. Fitur tambahan seperti dropdown list digunakan untuk mempermudah pemilihan data spesifik, seperti tingkat pendidikan. Seluruh proses ini menghasilkan dashboard interaktif yang dapat memvisualisasikan data dengan cara yang mudah dipahami oleh pengguna.

Setelah seluruh tahap pembuatan *dashboard* dilakukan dapat dihasilkan *Dashboard* Pengaruh Pendidikan Terhadap Pendapatan Individu yang dapat diakses pada link berikut <https://lookerstudio.google.com/s/sl8BW0THYeQ>. Dari dashboard ini, terdapat beberapa insight yang dapat diambil. Analisis menunjukkan bahwa rata-rata pendapatan individu sekitar Rp.3.598.788, dengan tingkat pendidikan S1/D4/D3 memiliki pendapatan tertinggi, yaitu sekitar Rp.5.500.000. Individu yang memiliki sertifikasi cenderung memiliki pendapatan lebih tinggi, dan laki-laki umumnya memiliki pendapatan lebih tinggi dibanding perempuan, terutama pada tingkat pendidikan

S1/D4/D3. Pekerjaan seperti dokter juga menunjukkan bahwa pendidikan berpengaruh terhadap pendapatan. Pada tabel ringkasan, individu wiraswasta dengan pendidikan SMA memiliki pendapatan tertinggi, sementara pada petani, lama bekerja berpengaruh terhadap pendapatan, sedangkan pada pedagang, usia dan lama bekerja tidak menunjukkan pengaruh signifikan terhadap pendapatan.



Gambar 12 Dashboard

PEMBAHASAN

1. Analisis ANOVA ONE WAY

Hasil analisis dengan menggunakan Anova One Way menunjukkan bahwa tingkat pendidikan berpengaruh terhadap pendapatan individu. Individu dengan tingkat pendidikan yang lebih tinggi cenderung memiliki pendapatan lebih besar dibandingkan dengan mereka yang memiliki tingkat pendidikan lebih rendah. Pendapatan individu tertinggi diraih oleh individu dengan tingkat pendidikan S1/D4/D3 dengan rata-rata pendapatan Rp. 5.500.000. Selain itu terdapat perbedaan yang signifikan antara pendapatan individu pada tingkat pendidikan S1/D4/D3 & SMA dan antara tingkat pendidikan S1/D4/D3 & SD ditandai dengan p-adj 0,00 lebih kecil dari taraf signifikansi 0,05 (5%) . Melalui analisis anova one way keahlian/sertifikasi berpengaruh pada pendapatan individu setiap tingkat pendidikan, baik SD, SMP, SMA, maupun S1/D4/D3 karena pada setiap tingkat pendidikan individu yang memiliki keahlian/sertifikasi dengan persentase 58,3% memiliki pendapatan lebih tinggi dibanding yang tidak memiliki keahlian/sertifikasi dengan persentase 41,7%. Pada hasil analisis melalui looker studio, individu yang berjenis kelamin laki-laki mempunyai rata-rata pendapatan lebih tinggi dibanding perempuan yang membuat jenis kelamin berpengaruh terhadap pendapatan individu. Pada Jenis pekerjaan, pekerjaan yang tingkat pendidikannya S1/D4/D3 memiliki pendapatan tertinggi. Selanjutnya, ada beberapa pekerjaan yang tidak dipengaruhi oleh tingkat pendidikan, yakni wiraswasta, petani, dan pedagang. Pengaruh lama bekerja dan usia tergantung kepada jenis pekerjaan, dalam penelitian ini, pada wiraswasta dan pedagang lama bekerja dan usia tidak mempengaruhi jumlah pendapatan, tetapi pada petani faktor lama bekerja

mempengaruhi jumlah pendapatan.

2. Dashboard Looker Studio

Dari dashboard ini terdapat beberapa insight sebagai berikut :

- Analisis scorecard. Pada scorecard yang dibuat terlihat bahwa rata-rata seluruh pendapatan

- individu pada data yang diolah yaitu sekitar Rp.3.598.788.
- b. Analisis column chart. Pada column chart didapat bahwa tingkat pendidikan S1/D4/D3 memiliki rata-rata pendapatan tertinggi diantara tingkat pendidikan lainnya yaitu sekitar Rp.5.500.000. Analisis donut chart. Pada donut chart dapat dilihat persentase rata-rata pendapatan individu yang memiliki sertifikasi lebih banyak daripada yang individu yang tidak memiliki sertifikasi. Lalu jika kita menggunakan scorecard, jika dilihat satu persatu didapat bahwa pada setiap tingkat pendidikan yang individu yang memiliki rata-rata pendapatan tertinggi mereka yang mempunyai keahlian/sertifikasi.
 - c. Analisis treemap chart. Pada treemap chart dapat dilihat bahwa individu yang berjenis kelamin laki-laki memiliki pendapatan yang lebih tinggi dibandingkan individu yang berjenis kelamin perempuan. Lalu juga didapat pada jenis kelamin laki-laki maupun perempuan juga didapat yang memiliki tingkat pendidikan S1/D4/D3 yang memiliki pendapatan tertinggi.
 - d. Analisis stacked bar chart. Pada stacked bar chart dapat dilihat pengaruh pendapatan berdasarkan jenis pekerjaan dan pendidikan, dimana terdapat pekerjaan dengan rata-rata pendapatan tertinggi yaitu dokter dengan tingkat pendidikan S1/D4/D3. Dimana 3 jenis pekerjaan teratas adalah pekerjaan dengan tingkat pendidikan S1/D4/D3. Maka dapat disimpulkan jenis pekerjaan dan tingkat pendidikan mempengaruhi pendapatan individu.
 - e. Analisis Tabel Ringkasan. Terakhir, pada tabel ringkasan individu, dapat dilihat detail mengenai masing-masing individu yang dianalisis. Melanjutkan analisis sebelumnya ada beberapa pekerjaan yang memiliki tingkat pendidikan yang bervariasi yaitu wiraswasta, petani, dan pedagang. Pada jenis pekerjaan wiraswasta, individu yang memiliki pendapatan tertinggi yaitu pada mereka yang berpendidikan SMA, jika pada stacked chart jenis pekerjaan wiraswasta yang berpendidikan SMA di klik, pada tabel ringkasan muncul 2 individu yang memiliki pendapatan yang tertinggi, pendapatan keduanya sama walaupun lama bekerja dan usia mereka berbeda, jadi dapat disimpulkan pada jenis pekerjaan wiraswasta lama bekerja dan usia tidak berpengaruh.

Lalu, pada jenis pekerjaan petani, individu yang memiliki pendapatan tertinggi yaitu pada mereka yang berpendidikan SMP, jika pada stacked chart di klik setiap tingkat pendidikan pada jenis pekerjaan petani, dapat dibandingkan yang berpendidikan SMP memiliki lama bekerja lebih lama dibandingkan individu pada tingkat pendidikan lainnya. Jadi dapat diambil kesimpulan pada jenis pekerjaan petani lama berkeja yang mempengaruhi pendapatan.

Selanjutnya pada jenis pekerjaan pedagang, individu yang memiliki pendapatan tertinggi yaitu pada mereka yang berpendidikan SMP, jika pada stacked chart jenis pekerjaan petani yang berpendidikan SMP di klik, pada tabel ringkasan muncul beberapa individu yang memiliki pendapatan yang tertinggi. Dapat dilihat bahwa lama bekerja dan usia tidak berpengaruh terhadap pendapatan, dengan membandingkan lama bekerja dan usia individu-individu tersebut.

KESIMPULAN

Penelitian ini mengeksplorasi keterkaitan antara tingkat pendidikan dan pendapatan individu melalui analisis *Anova One Way*. Hasil menunjukkan pengaruh signifikan pendidikan terhadap pendapatan, di mana individu berpendidikan S1/D4/D3 memiliki pendapatan lebih tinggi dibanding mereka yang berpendidikan SMA atau lebih rendah, dengan tingkat signifikansi yang kuat ($p\text{-adj} = 0,00$). Faktor penunjang sertifikasi atau keahlian juga berperan penting dalam meningkatkan pendapatan, dengan 58,3% individu bersertifikasi memperoleh pendapatan lebih tinggi dibandingkan yang tidak bersertifikasi. Melalui Looker Studio, dapat diamati bahwa individu berjenis kelamin laki-laki memiliki rata-rata pendapatan yang lebih tinggi dibandingkan dengan perempuan, menunjukkan bahwa jenis kelamin berpengaruh terhadap pendapatan individu. Berdasarkan jenis

pekerjaan, pekerjaan dengan tingkat pendidikan S1, D4, atau D3 cenderung menghasilkan pendapatan tertinggi. Namun, terdapat beberapa jenis pekerjaan yang tidak dipengaruhi oleh tingkat pendidikan, yaitu wiraswasta, petani, dan pedagang. Pengaruh lama bekerja dan usia terhadap pendapatan bergantung pada jenis pekerjaan. Dapat dilihat juga, pada wiraswasta dan pedagang, lama bekerja serta usia tidak memengaruhi jumlah pendapatan. Sebaliknya, pada petani, lama bekerja memiliki pengaruh terhadap jumlah pendapatan.

REFERENCES

- [1] M. B. Kapisa, S. A. Bauw, and R. A. Yap, "Analisis Tingkat Pendidikan dan Jenis Pekerjaan Terhadap Literasi Data dan Digital," vol. 03, no. 02, pp. 78–85, 2024.
Pendapatan Kepala Keluarga (KK) di Kampung Manbesak Distrik Biak Utara Provinsi Papua," *Lensa Ekon.*, vol. 15, no. 01, p. 131, 2021, doi: 10.30862/lensa.v15i01.145.
- [2] P. A. U. Dedi Julianto, "ANALISA PENGARUH TINGKAT PENDIDIKAN TERHADAP PENDAPATAN INDIVIDU DI SUMATERA BARAT Dedi," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.
- [3] P. Di and P. Banten, "Pengaruh tingkat pendidikan dan kesempatan kerja terhadap pengangguran di provinsi banten," 2012.
- [4] S. Hidayat, M. Saleh, U. B. Bangsa, K. Serang, K. Konseptual, and K. Sosial, "Komparasi kemampuan kerja antara lulusan sma dan smk di industri permesinan modern," vol. 2, no. 1, pp. 45–56.
- [5] Yeni Setiani, Nabila Rachmah, and Indra Purnama, "Visualisasi Data Malnutrisi Anak Di Asia Menggunakan Looker Studio Serta Analisis Data Dengan Metode ANOVA," *J. Ilm. Sist. Inf. dan Ilmu Komput.*, vol. 3, no. 3, pp. 188–212, 2023, doi: 10.55606/juisik.v3i3.701.
- [6] B. D. M. Chelsey Hill, Lanqing Du, Marina Johnson, "Comparing programming languages for data analytics: Accuracy of estimation in Python and R," *Data Mining and Knowledge Discovery*. [Online]. Available: https://www.researchgate.net/publication/377930782_Comparing_programming_languages_for_data_analytics_Accuracy_of_estimation_in_Python_and_R
- [7] J. Couto and D. D. Ruiz, "New Trends in Big Data Profiling," no. July, 2022, doi: 10.1007/978-3-031-10461-9.
- [8] I. Melissa and R. S. Oetama, "Analisis Data Pembayaran Kredit Nasabah Bank Menggunakan Metode Data Mining," vol. IV, no. 1, pp. 18–27, 2013.
- [9] A. Satu, J. One, and W. A. Y. Anova, "Statistika penelitian," no. April, 2020.
- [10] A. Rasio *et al.*, "SESUDAH COVID-19 SUBSECTOR FOOD AND BEVERAGE," vol. 12, no. 3, pp. 564–572, 2023.
- [11] A. Wardhana, "Uji Asumsi Klasik," no. July, 2024.
- [12] T. Prasetyo and A. I. Sabilah, "7 1234567," vol. 2, no. 6, pp. 775–785, 2024.
- [13] D. A. Alfarez, M. R. Ramadhan, P. Matematika, and U. Jambi, "Multi Proximity: Jurnal Statistika Universitas Jambi Anova dan Tukey HSD Perbandingan Produksi Padi Antara Tiga Kabupaten di Provinsi Anova and Tukey HSD Comparison of Rice Production Between Three Regencies in Jambi Padi merupakan komoditas tanaman yang telah lama dibudidayakan oleh petani , dimana untuk kegiatan dimana dibutuhkan faktor-faktor produksi yang di dalam ilmu ekonomi terdiri dari industri pertanian padi , antara lain kabupaten Merangin , Kerinci dan Tanjung Jabung Timur untuk membedakan nilai rata-rata lebih dari dua kelompok data dengan cara membandingkan," vol. 2, no. 1, pp. 23–31, 2023.
- [14] A. S. Wear, "Anova Satu Jalur (One Way ANOVA) dengan SPSS," *wordpress.com*. Accessed: Dec. 21, 2020. [Online]. Available: <https://alisadikinwear.wordpress.com/2020/12/21/anova-satu-jalur-one-way-anova-dengan-spss/>
- [15] I. Purnama, Y. Setiani, F. Ari, and N. Wibisono, "Analisis Dan Visualisasi Data Menggunakan Looker Studio Pada Dataset New York City Property Sales," vol. 13, pp. 2222–2234, 2025.
- [16] S. D. Puspa, J. Riyono, and C. E. Pujiastuti, "Pelatihan Pembuatan Dashboard dan Visualisasi Data dengan Looker Studio untuk Penguatan