*Journal of Student Development Information System (JoSDIS)*
**e-ISSN:2774-7948**
**Volume 2 Nomor 2, Juli 2022**
**Program Studi Sistem Informasi**
**Fakultas Sains dan Teknologi**
**Universitas Labuhanbatu**

# Penggunaan Bahasa Pemrograman Python Untuk Menganalisis Faktor Risiko Penyakit Diabetes Pada Masyarakat Keturunan Pima Indian

**[1]Steve Aditya Gandha, [2]Yonathan Immanuel Nurhan, [3]Jeffrey William**

[1,2,3]Computer Science & Mathematics, Universitas Bina Nusantara

Email : [1]steve.gandha@binus.ac.id, [2]yonathan.nurhan@binus.ac.id, [3]jeffrey.william@binus.ac.id

*Corresponding Mail Author:* steve.gandha@binus.ac.id

***Abstract***

*Diabetes is a disease that has existed for a long time. Until now, the prevalence of this disease is still quite high. Diabetes can occur in people of all ages, both children and adults. There are several factors that can increase the risk of diabetes either from genetic factors, diet, or age. In this study the authors will analyze these factors to see which factors have a strong correlation with diabetes. The analysis will be carried out using the Python programming language with the EDA (exploratory data analysis) method where the data will go through several stages starting from data collection, cleaning, to analysis. Analysis will be assisted by several libraries such as pandas, seaborn, & matplotlib. Data will be presented in tabular form and also in diagram form. In this study the author will also discuss the results of the python code. From this study it was found that things such as body weight, number of pregnancies, and the amount of glucose have a fairly high correlation with the risk of diabetes. Factors that are not natural factors can be prevented/maintained by doing a number of things such as exercising, maintaining food, etc.*

Keywords: *Python, Diabetes, Pandas, Seaborn, Matplotlib.*

## Pendahuluan

Masalah-masalah kesehatan yang ada di dunia saat ini dapat berupa penyakit yang menular maupun yang tidak menular, penyebabnya pun berbagai macam. Mulai dari faktor genetik, pola makan, pola tidur, dan masih banyak lagi. Salah satu dari penyakit itu adalah diabetes. Diabetes saat ini prevalensinya sedang meningkat baik pada kaum anak-anak, remaja, maupun dewasa. Menurut Kementerian Kesehatan, terjadi peningkatan yang cukup signifikan sebanyak 1.6% yaitu 6,9% pada tahun 2013 menjadi 8.5% pada tahun 2018. Beberapa provinsi yang memiliki peningkatan cukup signifikan adalah NTT (1.2% - 3.3%), Sulawesi Utara (0.8% - 3.4%), dan Maluku (0.5% - 2.1%) (Riskesdas, 2018). Tidak hanya pada negara berkembang saja, namun diabetes melitus juga memiliki prevalensi yang cukup tinggi pada negara-negara maju seperti Amerika (Azis et al., 2020). Diabetes tergolong sebagai sebuah penyakit degeneratif dan kronis. Penyakit degeneratif adalah penyakit kronis/jangka panjang yang dapat mempengaruhi pola hidup seseorang, dengan mengurangi tingkat aktivitas & tingkat produktivitas. Hal ini juga menyebabkan tingkat kematian yang tinggi pada

penyakit degeneratif (Hanum, Ardiansyah, 2018). Hal yang membuat penyakit ini sangat berbahaya adalah komplikasi yang disebabkannya. Beberapa diantaranya adalah stroke, penyempitan vaskuler, kegagalan fungsi organ jantung, mata, ginjal dan masih banyak lagi (Muliasari et al., 2019). Pengelolaan, pencegahan, dan deteksi dini diabetes sangatlah penting untuk dilakukan untuk mengurangi risiko terjadinya kematian. Hal-hal tersebut akan dianalisis menggunakan sebuah bahasa Pemrograman yaitu Python. Python diciptakan oleh Guido van Rossum di Belanda pada tahun 1990 dan namanya diambil dari acara televisi kesukaan Guido Monty yaitu *Python's Flying Circus.* (Mulki et al., 2022). Python memiliki banyak *built in libraries* yang dapat memudahkan user saat menggunakan bahasa tersebut, salah satunya adalah Pandas untuk mengolah tabel. Untuk membantu visualisasi data yang telah diolah menggunakan pandas, Python memiliki *built in libraries* lainnya seperti matplotlib, seaborn, dan masih banyak lagi. Pada *libraries* tersebut user dapat melihat hasil dari data yang telah diolah dengan pandas dalam bentuk grafik atau diagram.

Pada penelitian ini penulis akan melakukan analisis terhadap beberapa variabel untuk mencari korelasi mereka dengan penyakit diabetes Penelitian ini dilakukan dengan menganalisis *dataset* yang berasal dari *kaggle.* Hasil dari penggunaan bahasa Python dapat berbentuk tabel, graf, diagram, dll. Dari hasil-hasil tersebut kita akan dapat menentukan faktor-faktor apa saja yang dapat mempengaruhi risiko diabetes pada seseorang dan bagaimana cara mencegah atau menjaga agar penyakit tersebut tidak mengancam jiwa kita.

## Landasan Teori

### *Python*

*Python* merupakan salah satu bahasa Pemrograman terpopuler di dunia. Hal ini disebabkan karena Python dapat digunakan dalam berbagai macam bidang baik itu *Data Science, Artificial Intelligence, object oriented programming,* dan masih banyak lagi. Tidak hanya itu, Python juga merupakan sebuah bahasa Pemrograman yang sangat mudah untuk dimengerti karena *syntax* nya yang cukup simpel. (Mulki et al., 2022)

### *Python Libraries*

Dalam *Python* terdapat beberapa *built in libraries* yang dapat membantu penggunanya dalam membuat suatu program. *Libraries-libraries* ini berfungsi untuk membuat pengerjaan program menjadi lebih mudah. Pada *built in libraries* terdapat banyak fungsi yang telah ada sebelumnya dan sudah siap untuk dijalankan.

### Diabetes

Diabetes melitus merupakan sebuah penyakit tidak menular yang tidak dapat disembuhkan, sehingga penderita harus melakukan pengelolaan gula darah seumur hidup (Nasution et al., 2022)

### *EDA*

*EDA* merupakan salah satu cara menganalisis/visualisasi data yang dapat membantu dalam memahami data yang ada. Pada *EDA*, data akan dipelajari dan

dianalisis secara mendalam mulai dari membaca data, membersihkan data, menemukan *pattern* pada data dan juga menemukan korelasi pada data yang ada.(Radhi et al., 2022)

**Kaggle**

Kaggle adalah suatu komunitas online bagi para peneliti maupun peminat di bidang data science atau machine learning yang menyediakan kumpulan data, kompetisi, source code, tutorial, dan forum diskusi (Fonseca et al., 2018).

**Pandas**

Pandas merupakan salah satu *built in library* pada python yang telah dioptimalisasi untuk membantu user dalam mengolah data bentuk tabel (Lemenkova, 2019).

## Metode Penelitian

Populasi dari penelitian ini adalah para wanita keturunan *Pima Indian* yang berumur 21 tahun. Metode pengumpulan datanya sendiri berupa metode kuantitatif. Analisis yang akan dilakukan berupa *exploratory data analysis (EDA)* menggunakan bahasa Pemrograman *Python*.

Tahap pertama yang dilakukan adalah *data collection* atau membaca dan menyimpan data yang didapat dari *website* kaggle dalam bentuk *csv*. Pada tahap ini, *dataset* akan dibaca menggunakan fungsi dari *libraries-libraries* yang sudah dipanggil dan akan disimpan pada sebuah variabel. Setelah data disimpan dan dibaca, *size* dari tabel yang keluar akan diatur menggunakan fungsi dari *library* pandas.

```
#data collection
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

df = pd.read_csv("diabetes.csv")
desired_width = 320
pd.set_option('display.width', desired_width)
pd.set_option('display.max_columns',15)
pd.options.display.float_format = "{:.2f}".format
```

**Gambar 1. *Code* Untuk *Data Collection***

Untuk tahap selanjutnya adalah *data cleaning* atau proses membersihkan data dengan cara mengkoreksi, mengganti, atau menghapus data-data yang tidak diperlukan. data-data yang tidak diperlukan dapat disebabkan oleh beberapa faktor seperti data yang hilang, data yang ukurannya kecil, data yang duplikat atau data yang tidak konsisten. Pada *dataset* ini, proses pembersihan data yang dilakukan adalah menghilangkan data-data yang duplikat.

```
#data cleaning
duplicate_rows_data = df[df.duplicated()]
print("number of duplicate rows: ", duplicate_rows_data.shape)
for column in df.columns:
    num_distinct_values = len(df[column].unique())
    print(f"{column}: {num_distinct_values} distinct values")
print(df.isnull().sum())
print(df)
```

**Gambar 2. *Code* Untuk *Data Cleaning***

Setelah selesai melakukan *data cleaning*, tahap selanjutnya adalah melakukan *data visualization*. Bagian *data visualization* sendiri dibagi menjadi 2 yaitu *univariate analysis* dan *bivariate analysis.* Pada tahap *univariate analysis*, visualisasi/analisis akan dilakukan pada variabel secara individu. variabel-variabel yang ada akan dilihat distribusinya baik menggunakan *barplot*, *histogram*, ataupun *scatterplot*. Proses visualisasi selanjutnya yaitu *bivariate analysis*. Pada tahap ini, variabel *outcome* akan dibandingkan satu per satu dengan variabel-variabel lainnya secara numerik dan juga melalui visualisasi menggunakan *boxplot/scatterplot* untuk melihat korelasi mereka. Setelah kedua analysis itu dilakukan variabel outcome akan dibandingkan menggunakan heatmap dari *library* seaborn untuk melihat perbandingan variabel *outcome* dengan variabel-variabel lainnya

```
#univariate analysis
for col in ['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin', 'BMI', 'DiabetesPedigreeFunction', 'Age']:
    plt.hist(df[col], bins=25, edgecolor='black')
    plt.title(f'{col} Distribution')
    plt.xlabel(f'{col}')
    plt.ylabel('Count')
    plt.show()

sns.countplot(x='Outcome', data=df)
plt.title('Diabetes Outcome Distribution')
plt.show()
```

**Gambar 3. *Code* Untuk *Univariate Analysis***

```
#multivariate analysis
for col in ['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin', 'BMI', 'DiabetesPedigreeFunction', 'Age']:
    print(df[['Outcome', col]].corr(), "\n")
    sns.boxplot(x='Outcome',  y=col, data=df)
    plt.title(f'{col} vs Diabetes')
    plt.xlabel(f'Diabetes')
    plt.ylabel(col)
    plt.show()

sns.set(font_scale=0.8)
sns.set_style("white")
sns.set_palette("PuBuGn_d")
plt.title('Correlation with Diabetes')
sns.heatmap(df.corr()[['Outcome']].sort_values(by='Outcome', ascending=False), annot=True)
plt.show()
```

**Gambar 4. *Code* Untuk *Multivariate Analysis***

## Hasil Dan Pembahasan

| | Pregnancies | Glucose | BloodPressure | SkinThickness | Insulin | BMI | DiabetesPedigreeFunction | Age | Outcome |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 6 | 148 | 72 | 35 | 0 | 33.60 | 0.63 | 50 | 1 |
| 1 | 1 | 85 | 66 | 29 | 0 | 26.60 | 0.35 | 31 | 0 |
| 2 | 8 | 183 | 64 | 0 | 0 | 23.30 | 0.67 | 32 | 1 |
| 3 | 1 | 89 | 66 | 23 | 94 | 28.10 | 0.17 | 21 | 0 |
| 4 | 0 | 137 | 40 | 35 | 168 | 43.10 | 2.29 | 33 | 1 |
| .. | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 763 | 10 | 101 | 76 | 48 | 180 | 32.90 | 0.17 | 63 | 0 |
| 764 | 2 | 122 | 70 | 27 | 0 | 36.80 | 0.34 | 27 | 0 |
| 765 | 5 | 121 | 72 | 23 | 112 | 26.20 | 0.24 | 30 | 0 |
| 766 | 1 | 126 | 60 | 0 | 0 | 30.10 | 0.35 | 47 | 1 |
| 767 | 1 | 93 | 70 | 31 | 0 | 30.40 | 0.32 | 23 | 0 |

[768 rows x 9 columns]

**Gambar 5. Hasil *Print* Data Yang Didapat Dari *Kaggle***

```
number of duplicate rows:  (0, 9)

Pregnancies: 17 distinct values
Glucose: 136 distinct values
BloodPressure: 47 distinct values
SkinThickness: 51 distinct values
Insulin: 186 distinct values
BMI: 248 distinct values
DiabetesPedigreeFunction: 517 distinct values
Age: 52 distinct values
Outcome: 2 distinct values
```

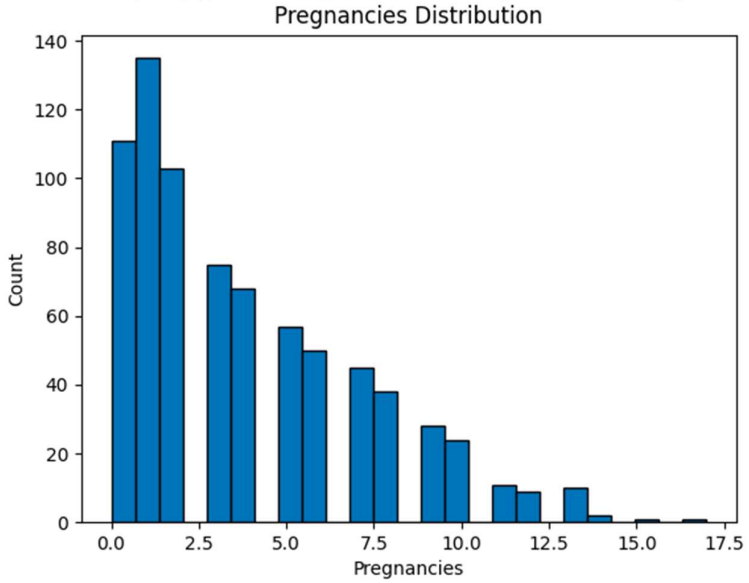
**Gambar 6. Banyaknya Duplikat Dan *Value-value* Unik Pada Tiap Variabel**

Dapat dilihat bahwa terdapat *duplicate rows* sebanyak 9 *rows,* sehingga kesembilan data duplikat tersebut akan dihapus pada tahap *data cleaning*.

```
 Pregnancies                 0
Glucose                      0
BloodPressure                0
SkinThickness                0
Insulin                      0
BMI                          0
DiabetesPedigreeFunction     0
Age                          0
Outcome                      0
```

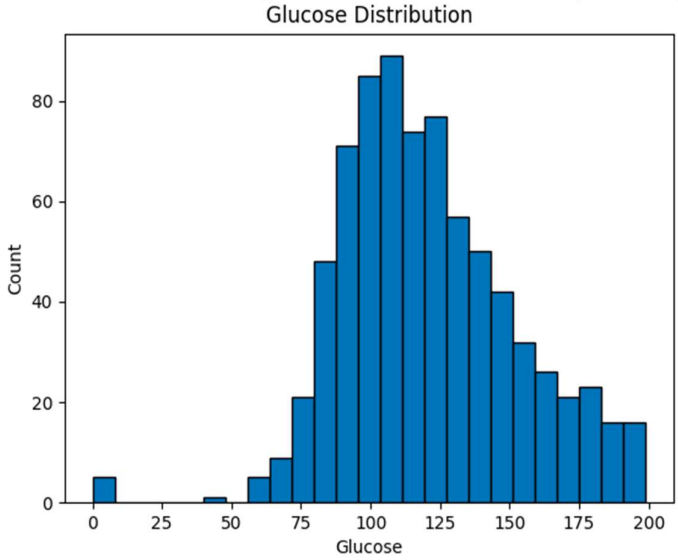
**Gambar 7. Banyaknya *NULL Value* Pada Tiap Variabel**

Pada *dataset* tersebut, tidak terdapat data *NULL* pada variabel-variabel yang ada sehingga tidak ada data yang perlu dihapus dalam tahap *data cleaning*

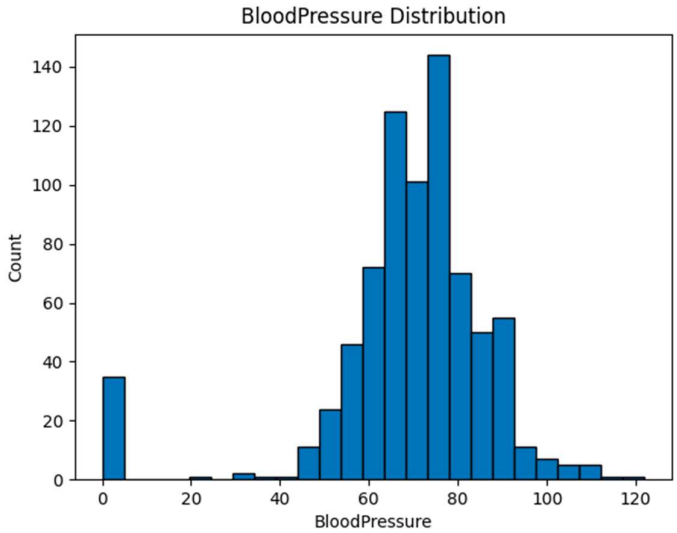



**Gambar 8. Histogram Penyebaran *Pregnancies***

*Pregnancies* pada data ini merupakan jumlah kehamilan dari tiap responden. Pada *dataset* ini sebagian besar orang mengalami kehamilan sebanyak 1-2 kali dan dapat dilihat semakin besar *value Pregnancies*, semakin sedikit jumlahnya.

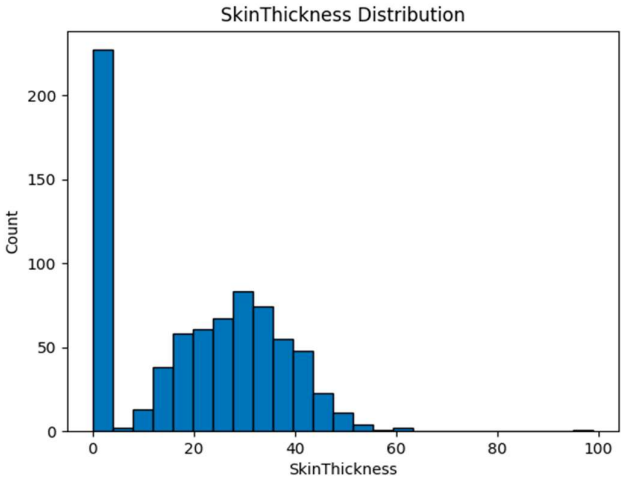



**Gambar 9. Histogram Penyebaran *Glucose***

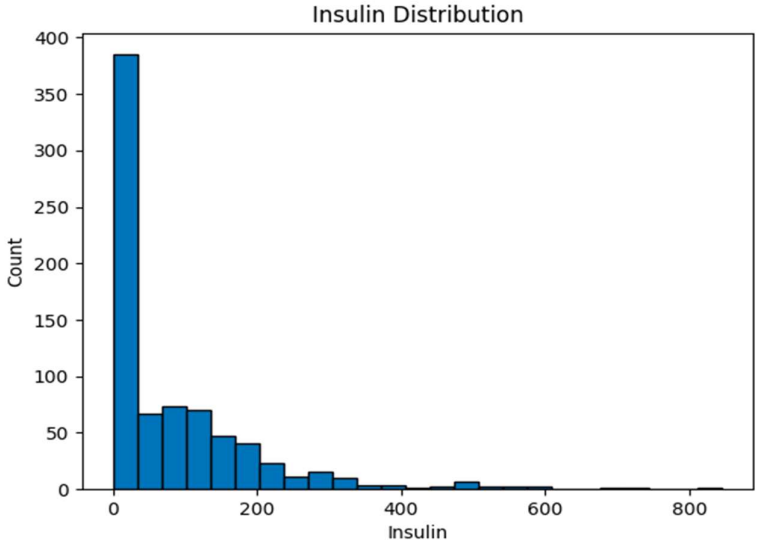



**Gambar 10. Histogram Penyebaran *Blood Pressure***

*Glucose* pada *dataset* ini adalah tingkat *plasma glucose* dalam darah yang didapat dari tes. *BloodPressure* dalam *dataset* ini adalah tingkat tekanan darah dari tiap responden(mm/Hg). Pada Histogram *Glucose* dan *Blood Pressure* dapat dilihat penyebaran data cukup merata, namun terdapat *outliers* pada kedua variabel tersebut. Pada kedua *dataset* ini, *outliers* tersebut akan dihapus karena *value* tersebut tidak logis.

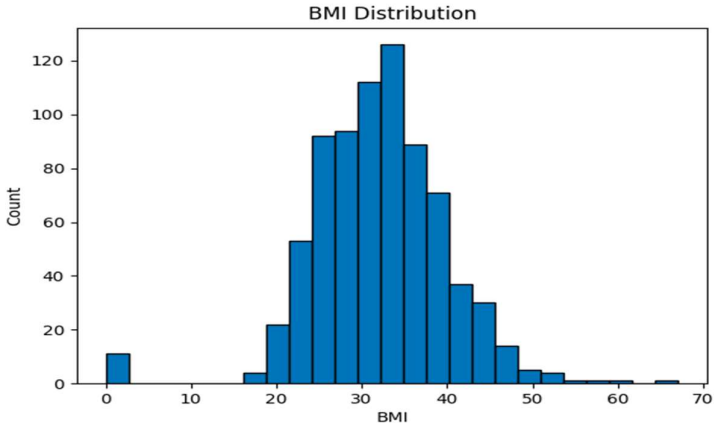



**Gambar 11. Histogram Penyebaran *Skinthickness***



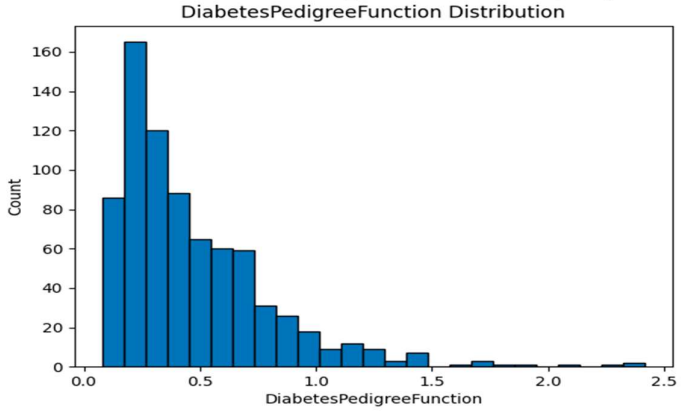

**Gambar 12. Histogram Penyebaran *Insulin***

*SkinThickness* pada *dataset* ini adalah ketebalan kulit dalam mm, angka yang lebih tinggi dapat mengindikasikan obesitas. *Insulin* adalah *2-hour serum insulin* dalam (mu U/ml). Pada histogram tersebut, *SkinThickness* dan *Insulin* memiliki distribusi data yang *skewed* ke kiri atau bisa dibilang sebagian besar responden memiliki tingkat *Skin Thickness & Insulin* yang rendah.

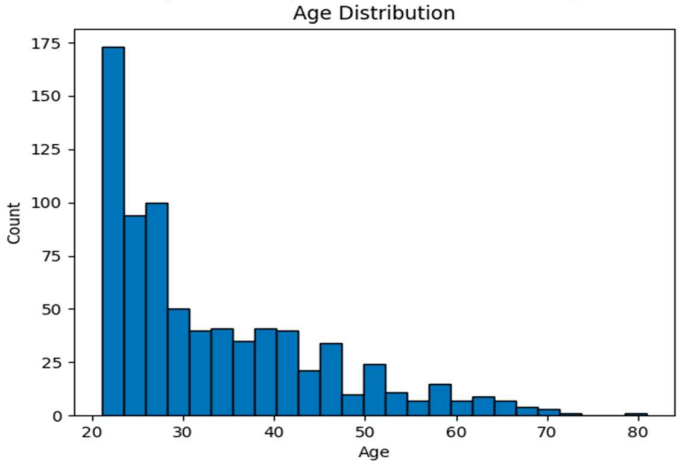



**Gambar 13. Histogram penyebaran *BMI***

*BMI* pada *dataset* ini adalah *Body Mass Index* dari tiap responden yang dihitung dengan rumus (Berat/Tinggi^2). Variabel ini memiliki distribusi yang merata dengan sedikit *outliers*. *Outliers* tersebut akan dihapus karena *value* yang tidak logis.

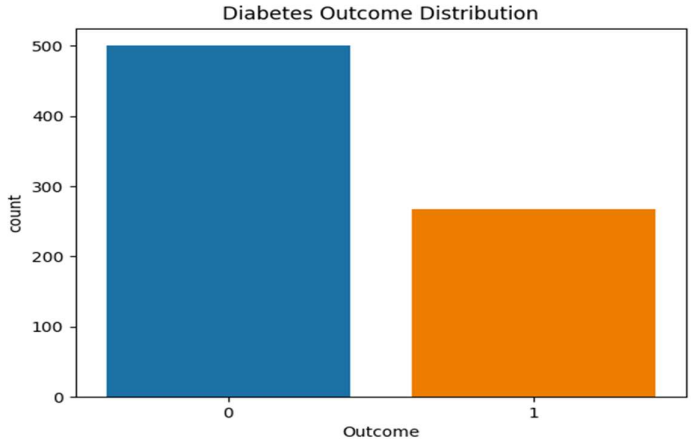



**Gambar 14. Histogram Penyebaran *Diabetes Pedigree Function***



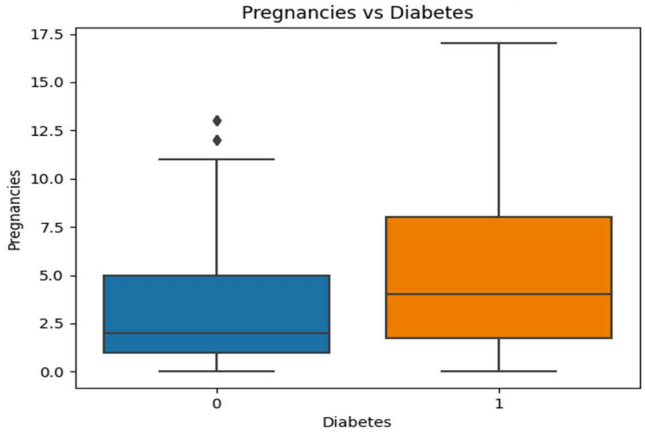

**Gambar 15. Histogram Penyebaran *Age***

*Diabetes Pedigree Function* merupakan resiko seseorang mendapatkan diabetes dari genetik keluarga, sedangkan *Age* dalam *dataset* ini adalah umur dari tiap responden. Pada Histogram penyebaran *Age* dan *Diabetes Pedigree Function*, kedua *dataset* tersebut memiliki distribusi data yang *left skewed*. Yang berarti sebagian besar responden memiliki resiko genetik yang rendah & berusia dibawah 30 tahun.




**Gambar 16. *Countplot* Penyebaran *Outcome***

*Outcome* pada *dataset* memiliki 2 buah *value* yaitu 0(*no*) atau 1(*yes*). 0 berarti responden tidak memiliki diabetes, sedangkan 1 berarti responden memiliki diabetes.

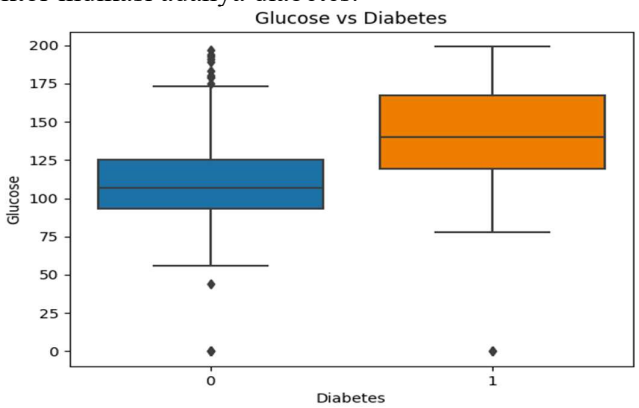



**Gambar 17. *Boxplot Pregnancies* vs *Diabetes***

**Tabel 1. *Outcome and Pregnancies***

| | *Outcome* | *Pregnancies* |
|---|---|---|
| *Outcome* | 1.00 | 0.22 |
| *Pregnancies* | 0.22 | 1.00 |

Dapat dilihat dari *boxplot* tersebut, responden yang memiliki diabetes memiliki *spread pregnancies* yang lebih tinggi. Dapat dilihat juga bahwa *pregnancies* memiliki korelasi yang cukup *moderate* dengan diabetes yaitu sebesar 0.22. Dapat diartikan bahwa jumlah kehamilan yang tinggi dapat menjadi salah satu faktor penyebab diabetes atau menjadi faktor indikasi adanya diabetes.

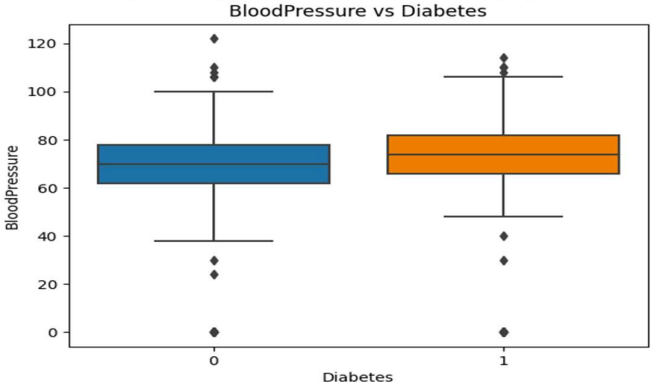



**Gambar 18.** ***Boxplot Glucose vs Diabetes***

**Tabel 2,** ***Outcome and Glucose***

| | . | *Glucose* |
|---|---|---|
| *Outcome* | 1.00 | 0.47 |
| *Glucose* | 0.47 | 1.00 |

Dapat dilihat dari *boxplot* tersebut, responden yang memiliki diabetes memiliki *spread glucose* yang jauh lebih tinggi dibandingkan dengan responden tanpa diabetes. Dapat dilihat juga bahwa *glucose* memiliki korelasi yang tinggi dengan diabetes yaitu sebesar 0.47. Variabel ini memiliki tingkat korelasi dengan Diabetes yang paling tinggi. Dimana dapat diartikan bahwa tingkat *glucose* yang tinggi dapat menjadi faktor kuat penyebab diabetes atau dapat menjadi indikasi kuat adanya penyakit diabetes.

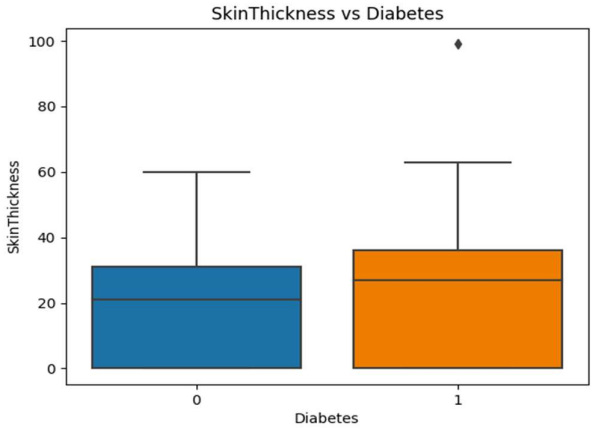



**Gambar 19.** ***Boxplot Blood Pressure vs Diabetes***

**Tabel 3. *Outcome and Blood Pressure***

| | *Outcome* | *Blood Pressure* |
|---|---|---|
| *Outcome* | 1.00 | 0.065 |
| *BloodPressure* | 0.065 | 1.00 |

Dapat dilihat dari *boxplot* tersebut, responden yang memiliki diabetes memiliki *spread Blood Pressure* yang hampir sama dengan responden tanpa diabetes. Dapat dilihat juga bahwa *Blood Pressure* memiliki korelasi yang cukup rendah dengan diabetes yaitu hanya sebesar 0.065.Variabel ini memiliki tingkat korelasi dengan Diabetes yang paling rendah. Dari hal-hal tersebut dapat diartikan bahwa tekanan darah seseorang bukan merupakan faktor penyebab diabetes dan tidak dapat dijadikan indikasi adanya penyakit diabetes.

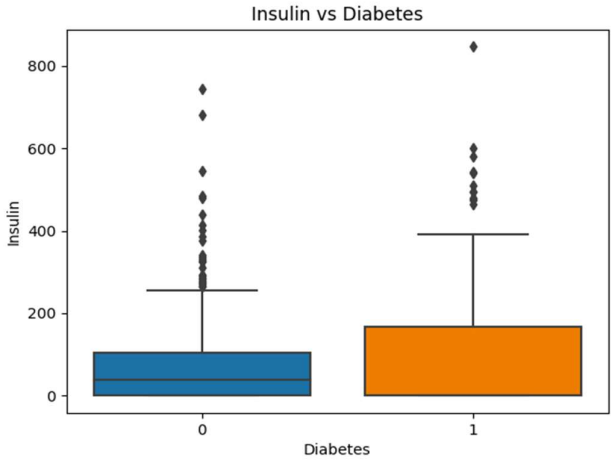



**Gambar 20. *Boxplot Skin Thickness vs Diabetes***

**Tabel 4. *Outcome and Skin Thickness***

| | *Outcome* | *Skin Thickness* |
|---|---|---|
| *Outcome* | 1.00 | 0.075 |
| *SkinThickness* | 0.075 | 1.00 |

Sama dengan *Blood Pressure*, responden yang memiliki diabetes memiliki *spread Skin Thickness* yang hampir sama dengan responden tanpa diabetes. Dapat dilihat juga bahwa *Skin Thickness* memiliki korelasi yang cukup rendah dengan diabetes yaitu hanya sebesar 0.075.Variabel ini memiliki tingkat korelasi dengan Diabetes yang hampir sama dengan tingkat korelasi *Blood Pressure* yang berarti *Skin Thickness* bukan merupakan faktor penyebab diabetes.

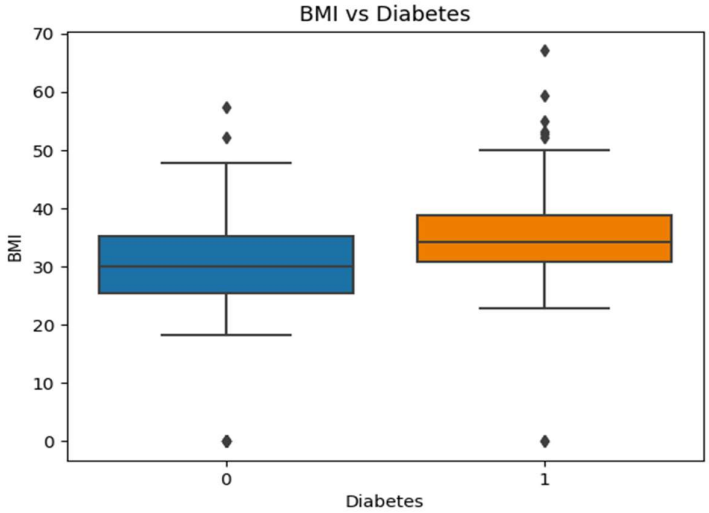



**Gambar 21. *Boxplot Insulin vs Diabetes***

**Tabel 5. *Outcome and Insulin***

| | *Outcome* | *Insulin* |
|---|---|---|
| *Outcome* | 1.00 | 0.13 |
| *Insulin* | 0.13 | 1.00 |

Dapat dilihat dari *boxplot* tersebut, responden yang memiliki diabetes memiliki *spread Insulin* yang sedikit lebih tinggi. Dapat dilihat juga bahwa *Insulin* memiliki korelasi yang cukup rendah atau tidak terlalu kuat dengan diabetes yaitu hanya sebesar 0.13. Dari kedua hal tersebut, dapat diartikan bahwa insulin dapat menjadi salah satu faktor penyebab diabetes, namun bukan merupakan faktor yang kuat.

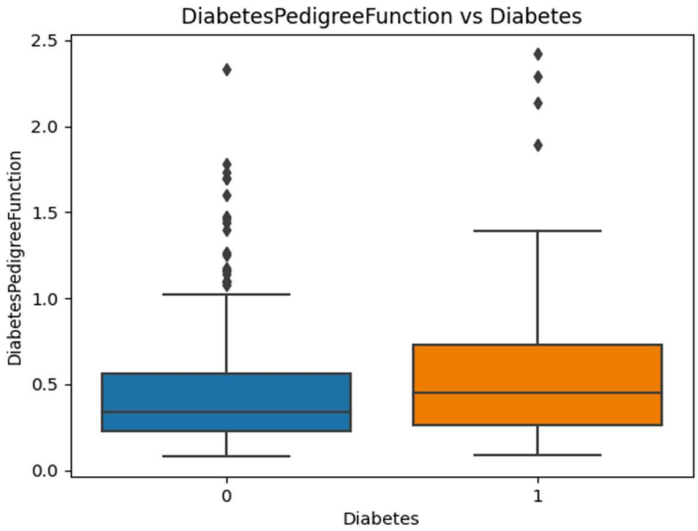



**Gambar 22. *Boxplot BMI vs Diabetes***

**Tabel 6. *Outcome and BMI***

| | . *Outcome* | *BMI* |
|---|---|---|
| *Outcome* | 1.00 | 0.29 |
| *BMI* | 0.29 | 1.00 |

Dapat dilihat dari *boxplot* tersebut, responden yang memiliki diabetes memiliki *spread BMI* yang lebih tinggi. Dapat dilihat juga bahwa *BMI* memiliki korelasi yang cukup *moderate* dengan diabetes yaitu sebesar 0.29. Dari kedua hal tersebut dapat diartikan bahwa *BMI* yang tinggi dapat mengindikasikan adanya penyakit diabetes dan menjadi faktor penyebab penyakit diabetes.

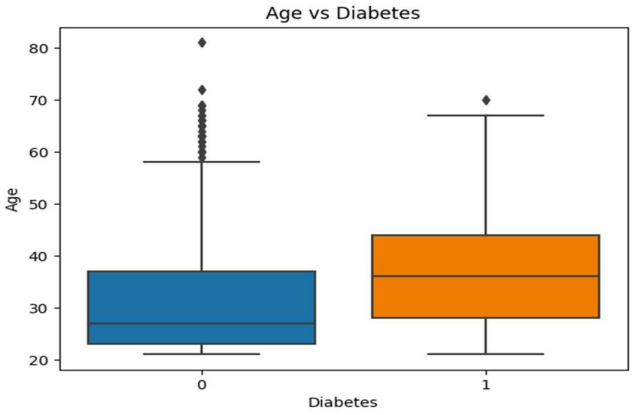



**Gambar 23. *Boxplot Diabetes Pedigree Function vs Diabetes***

**Tabel 7. *Outcome and Diabetes Pedigree***

| | *Outcome* | *Diabetes Pedigree* |
|---|---|---|
| *Outcome* | 1.00 | 0.17 |
| *Diabetes Pedigree* | 0.17 | 1.00 |

Dapat dilihat dari *boxplot* tersebut, responden yang memiliki diabetes memiliki *spread Diabetes Pedigree Function* yang sedikit lebih tinggi. Dapat dilihat juga bahwa *Diabetes Pedigree Function* memiliki korelasi yang cukup rendah atau tidak terlalu kuat dengan diabetes yaitu hanya sebesar 0.17. Dari kedua hal tersebut dapat diartikan bahwa faktor genetik dapat menjadi indikasi adanya penyakit diabetes, namun bukan merupakan indikasi yang kuat.




**Gambar 24. *Boxplot Age vs Diabetes***

**Tabel 8. *Outcome and Age***

| | *Outcome* | *Age* |
|---|---|---|
| *Outcome* | 1.00 | 0.24 |
| *Age* | 0.24 | 1.00 |

Dapat dilihat dari *boxplot* tersebut, responden yang memiliki diabetes memiliki *spread Age* yang lebih tinggi. Dapat dilihat juga bahwa *Age* memiliki korelasi yang mirip dengan BMI yaitu sebesar 0.24. Dari hal tersebut dapat diartikan bahwa usia yang lebih tua dapat menjadi faktor penyebab penyakit diabetes atau merupakan faktor indikasi yang cukup kuat.




**Gambar 25. *Heatmap* Korelasi Dengan Diabetes**

Dari *heatmap* tersebut dapat dilihat bahwa variabel yang memiliki korelasi paling tinggi dengan Diabetes adalah *Glucos*e sebesar 0.47 sampai yang terkecil yaitu *BloodPressure* sebesar 0.065.

## Kesimpulan

Dari hasil-hasil yang ada tersebut, dapat dilihat bahwa variabel yang memiliki korelasi tertinggi dengan resiko diabetes seseorang adalah tingkat glukosa dalam darah, diikuti oleh *BMI*, usia, dan jumlah kehamilan. Hal-hal tersebut juga dapat dijadikan variabel untuk memprediksi diabetes pada seseorang dan dapat menjadi faktor yang mempengaruhi resiko diabetes pada seseorang. Sedangkan hal seperti faktor genetik dan tingkat insulin memiliki korelasi yang *moderate* dengan diabetes, yang berarti variabel tersebut dapat digunakan untuk prediksi, namun tidak seakurat variabel seperti *BMI*, glukosa, dll. sedangkan variabel seperti tekanan darah & ketebalan kulit memiliki korelasi yang rendah dengan diabetes, yang berarti kedua variabel tersebut tidak dapat digunakan untuk melakukan prediksi dan kemungkinan besar bukan merupakan faktor penyebab diabetes pada seseorang. Beberapa hal seperti usia & faktor genetik merupakan hal yang tidak dapat dihindari/dicegah. Namun, hal lainnya yang memiliki korelasi tinggi seperti *BMI*, tingkat glukosa, atau jumlah kehamilan dapat dikendalikan. *BMI* dapat dikendalikan atau dijaga dengan olahraga yang cukup, tidak makan yang berlebihan, dan juga menjaga tingkat stress. Tingkat glukosa dalam darah dapat dijaga atau dicegah agar tidak terlalu tinggi dengan cara mengurangi mengkonsumsi makanan yang memiliki kadar gula yang tinggi dan juga tidur yang cukup. Sementara jumlah kehamilan dapat dikendalikan dengan keluarga berencana.

## Daftar Pustaka

Azis, W. A., Muriman, L. Y., & Burhan, S. R. (2020). Hubungan Tingkat Pengetahuan dengan Gaya Hidup Penderita Diabetes Mellitus. *Jurnal Penelitian Perawat Profesional*, 2(1), 105–114. https://doi.org/10.37287/jppp.v2i1.52

Fonseca, E., Plakal, M., Font, F., Ellis, D. P., Favory, X., Pons, J., & Serra, X. (2018). General-purpose tagging of freesound audio with audioset labels: Task description, dataset, and baseline. *arXiv:1807.09902*.https://doi.org/10.48550/arXiv.1807.09902

Hanum, G. R., & Ardiansyah, S. (2018). DETEKSI DINI PENYAKIT DEGENERATIF PADA REMAJA ANGGOTA KARANG TARUNA. *Jurnal Abadimas Adi Buana*, 2(1), 1–3. https://doi.org/10.36456/abadimas.v2.i1.a1615

Kemenkes, RI. (2018). Infodatin Diabetes. Jakarta: Pusat data dan Informasi Kemenkes

Lemenkova, P. (2019). PROCESSING OCEANOGRAPHIC DATA BY PYTHON LIBRARIES NUMPY, SCIPY AND PANDAS. *Aquatic Research*, 73–91. https://doi.org/10.3153/ar19009

Muliasari, H., Hamdin, C. D., Ananto, A. D., & Ihsan, M. (2019). EDUKASI DAN DETEKSI DINI DIABETES MELLITUS SEBAGAI UPAYA MENGURANGI PREVALENSI DAN RESIKO PENYAKIT DEGENERATIF. *Jurnal Pendidikan Dan Pengabdian Masyarakat*, 2(1). https://doi.org/10.29303/jppm.v2i1.1018

Mulki, A., Suhaedi, D., & Permanasari, Y. (2022). Optimasi Jaringan Distribusi Listrik dengan Pohon Rentang Minimum Menggunakan Bahasa Pemrograman Python. *Bandung Conference Series: Mathematics*, 2(1). https://doi.org/10.29313/bcsm.v2i1.1542.

Nasution, M. A., Nia Novranda Pertiwi, & Hindri Syahputri. (2022). Penyuluhan Penyakit Diabetes Melitus Dan Pemeriksaan Kesehatan Di Dusun I Desa Kolam Kecamatan Percut Sei Tuan. *Jurnal Pengabdian UntukMu NegeRI*, 6(2), 213–218. https://doi.org/10.37859/jpumri.v6i2.4017.

Radhi, M., Amalia, A., Sitompul, D. R. H., Sinurat, S. H., & Indra, E. (2022). ANALISIS BIG DATA DENGAN METODE EXPLORATORY DATA ANALYSIS (EDA) DAN METODE VISUALISASI MENGGUNAKAN JUPYTER NOTEBOOK. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, 4(2), 23–27. https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2475.