

Implementasi Data Analytics Guna Optimalisasi Kegiatan Penerimaan Santri Baru (PSB) Berbasis Dashboard

Diki Nurul Rivai*¹, Teguh Ikhlas Ramadhan*², Agus Supriatman*³

^{1,2,3}Universitas Perjuangan Tasikmalaya, Jl. Peta No.177,(0265) 326058

^{1,2,3}Teknik Informatika, Fakultas Teknik, Universitas Perjuangan Tasikmalaya

e-mail: *¹2003010050@unper.ac.id, ² teguhikhlas@unper.ac.id, agussupriatman@unper.ac.id

ABSTRAK

Penerapan data analytics di Pondok Pesantren Riyadlul Ulum Wadda'wah Condong dengan metode pengumpulan data, preprocessing, exploratory data analysis, evaluasi, dan visualisasi, memberikan kontribusi signifikan dalam optimalisasi kegiatan Penerimaan Santri Baru (PSB). Melalui penggunaan metode tersebut, analisis data dapat menghasilkan dashboard interaktif. Selain itu, analisis data juga memberikan wawasan dari pendaftaran santri baru, memberikan rekomendasi untuk meningkatkan efisiensi kegiatan PSB. Pendekatan komprehensif ini membantu memberikan panduan berbasis data bagi para pemangku kepentingan dalam pengambilan keputusan, mendukung implementasi kebijakan, dan merancang strategi yang dapat diukur guna optimalisasi kegiatan PSB.

Kata Kunci: Big data, Data Analis, Dashboard, Python, Looker Studio.

ABSTRACT

The implementation of data analytics at Pondok Pesantren Riyadlul Ulum Wadda'wah Condong, utilizing methods such as data collection, preprocessing, exploratory data analysis, evaluation, and visualization, significantly contributes to optimizing the New Student Admission (PSB) process. Through the application of these methods, data analysis can generate an interactive dashboard. Furthermore, the analysis provides insights into new student registrations, offering recommendations to enhance the efficiency of the PSB process. This comprehensive approach helps provide data-driven guidance for stakeholders in decision-making, supports policy implementation, and designs measurable strategies for optimizing PSB activities.

Keywords: Big Data, Data Analytics, Dashboard, Python, Looker Studio.

1. PENDAHULUAN

Dalam era digital yang berkembang pesat, data telah menjadi salah satu aset terbesar bagi berbagai instansi, termasuk Pondok Pesantren Riyadlul Ulum Wadda'wah Condong. Jumlah data yang terus meningkat secara signifikan, terutama dengan adanya Big Data yang mencakup data dalam jumlah besar dan kompleks. Memanfaatkan Big Data menjadi kunci strategis dalam pengambilan keputusan untuk kesuksesan sebuah organisasi[1]. Meskipun Pondok Pesantren Riyadlul Ulum Wadda'wah Condong memiliki jumlah data yang besar, sayangnya belum dapat memanfaatkannya sepenuhnya. Data tersebut mencakup informasi pendaftaran, data santri, data guru, dan aspek lainnya. Analisis dan pengolahan data dapat membantu instansi ini dalam mengoptimalkan program kerja dan pengambilan keputusan.

Sebagai contoh, analisis data pendaftaran dapat memberikan wawasan yang informatif untuk meningkatkan efisiensi kegiatan penerimaan santri baru (PSB).

Kegiatan PSB seringkali tidak optimal karena adanya pemborosan sumber daya, seperti penyebaran pamflet di lokasi yang kurang tepat dan kurang pemahaman terhadap target audience. Oleh karena itu, diperlukan analisis data pendaftaran untuk mengidentifikasi pola dan tren yang dapat membantu dalam pengambilan keputusan yang tepat. Tanpa analisis data yang tepat, pemecahan masalah dan pengambilan keputusan menjadi sulit [2]. Oleh karena itu, analisis data pendaftaran menjadi krusial untuk memberikan wawasan informatif. Manusia cenderung lebih mudah memahami informasi yang disajikan secara visual, sehingga penggunaan dashboard system sebagai alat visualisasi data dapat sangat membantu. Dashboard system menyajikan informasi penting menggunakan grafik, gambar, tabel, dan lainnya, memudahkan pemahaman data dengan cepat [3].

Analisis dan visualisasi data berbasis dashboard system diharapkan dapat memberikan kontribusi signifikan dalam optimalisasi kegiatan PSB di Pondok Pesantren Riyadlul Ulum Wadda'wah. Oleh karena itu, penelitian dan implementasi dashboard system menjadi relevan dalam konteks pengembangan strategi dan pengambilan keputusan yang efektif.

2. METODE PENELITIAN

2.1 Metodologi

Analisis dan visualisasi data pendaftaran santri berbasis dashboard system terdapat beberapa tahapan di dalamnya, dimulai dari pengumpulan data, proses data, analisis data, menafsirkan data dan visualisasi data [4]. Adapun untuk tahapannya seperti pada gambar 2.1 dibawah ini :



Gambar 1 Metode Penelitian

2.2 Data Collecting & Understanding

Pengumpulan data adalah proses pengoleksian informasi, fakta, atau elemen-elemen lainnya yang digunakan untuk tujuan analisis. Pada penelitian ini, tahapan pengumpulan data dilakukan dengan tiga cara yaitu observasi untuk memahami proses, selanjutnya wawancara untuk mendapatkan pendapat dan yang terakhir studi literatur dilakukan dengan mempelajari dan memahami literatur dari berbagai sumber.

2.3 Data Preprocessing

Dalam Data Preprocessing, langkah-langkah krusial dilakukan untuk persiapan data yang baik. Proses dimulai dengan Removing Duplicate untuk menciptakan dataset bersih. Data Cleaning melibatkan penambahan/penghapusan baris, termasuk normalisasi dan penanganan nilai yang hilang. Data Transformation digunakan untuk merubah format data agar sesuai dengan analisis. Pada tahap Data Reduction, dimensi data dikurangi dengan ekstraksi atau seleksi fitur. Tahapan ini meningkatkan validitas dan keakuratan hasil analisis [5].

2.4 *Exploratory Data Analysis*

Exploratory Data Analysis (EDA) adalah proses dalam data analytics untuk mengenal dan memahami dataset, mencari pola, trend, anomali, dan hubungan antar variabel. Tahapan EDA melibatkan penghitungan Central Tendency, penanganan Missing Value, eliminasi Outlier, dan interpretasi data[6].

2.5 *Evaluation*

Evaluasi merupakan proses penilaian kualitas analisis berdasarkan metrik-metrik yang telah digunakan. Fokusnya pada validasi hubungan korelasi variabel sebelum dan setelah preprocessing untuk menilai kehandalan informasi sebelum digunakan dalam perancangan strategi atau kebijakan. Penggunaan *Cramers V* sebagai langkah krusial dalam mengevaluasi elektabilitas variabel memastikan data dapat dipercaya dan berkontribusi signifikan dalam merancang langkah-langkah strategis yang efektif[7]. Validasi dan evaluasi bersama-sama memastikan analisis sesuai dengan tujuan dan memberikan gambaran tingkat keberhasilan analisis.

2.6 *Visualization*

Data visualisasi adalah seni mengubah data kompleks menjadi grafik, tabel, dan visualisasi yang mudah dipahami, dengan tujuan menyajikan informasi secara menarik, jelas, dan informatif. Alat seperti Looker Studio memainkan peran kunci dalam mencapai tujuan ini, memungkinkan pembuatan dashboard data interaktif dan terintegrasi. Dengan data visualisasi, kita dapat membandingkan data dan mengamati hubungannya, sementara manfaat utamanya mencakup peningkatan transparansi, pengurangan waktu analisis, dan pemahaman data yang lebih baik oleh pemangku kepentingan. Hal ini mendukung pengambilan keputusan yang lebih cerdas dan tepat[8].

3. HASIL DAN PEMBAHASAN

3.1 *Data Collecting dan Understanding*

Setelah melakukan Observasi, Wawancara dan studi literatur maka diperoleh dataset pendaftaran untuk di analisis dan diolah lebih lanjut. Pada sub bab ini, menjelaskan mengenai sekumpulan data yang telah diperoleh guna membangun dashboard system pada Pondok Pesantren Riyadlul Ulum Wadda'wah Condong. Sumber data diperoleh dari website resmi penerimaan santri baru Pondok Pesantren Riyadlul Ulum Wadda'wah Condong (<https://ppdb.pesantren-condong.net/>). Data tersebut berupa data pendaftaran santri baru dalam format Xlsx. Data yang terkumpulkan yaitu 5000 data terhitung dari tahun 2019 sampai dengan 2023. Dalam dataset pendaftaran santri baru, terdapat variabel-variabel didalamnya. Variabel tersebut bisa jadi acuan dalam pembangunan dashboard system. Adapun variabel di dalam dataset pendaftaran santri baru, sebagai berikut:

Tabel 1 Variabel Dataset Pendaftaran

No	Nama Variabel	Keterangan
1	Nama Lengkap	Informasi mengenai nama lengkap peserta
2	ID PPDB	No Unik peserta pendaftaran santri baru
3	Thn Ajaran	Informasi mengenai tahun ajaran yang didaftarkan peserta
4	Gelombang	Informasi mengenai gelombang pendaftaran santri
5	Jenjang	Informasi mengenai minat studi yang diinginkan calon santri (SMA/SMP)
6	Jenis daftar	Informasi mengenai peserta didik baru atau peserta didik lanjutan
7	Jenis Kelamin	Informasi mengenai jenis kelamin santri
8	NISN	Informasi mengenai nomor induk santri nasional
9	NIK	Informasi mengenai nomor induk kependudukan
10	No Reg Akta	Informasi mengenai nomor registrasi akta kelahiran calon santri
11	Berkebutuhan Khusus	Informasi mengenai memiliki atau tidaknya karakteristik khusus
12	Alamat	Informasi mengenai alamat calon santri
13	Rt	Informasi mengenai nomor rt calon santri
14	Rw	Informasi mengenai nomor rw calon santri
15	Dusun	Informasi mengenai dusun calon santri
16	Kelurahan/Desa	Informasi mengenai kelurahan atau desa calon santri
17	Kecamatan	Informasi mengenai kecamatan calon santri
18	Kabupaten/Kota	Informasi mengenai kabupaten atau kota calon santri
19	Provinsi	Informasi mengenai provinsi calon santri
20	Tempat Tinggal	Informasi mengenai tempat tinggal sebelum mondok(bersama orangtua/pesantren)
21	Anak Ke	Informasi mengenai info anak keberapa
22	No.KKS	Informasi mengenai nomor kartu keluarga sejahtera
23	KPS	Informasi mengenai memiliki atau tidaknya kartu perlindungan sosial
24	No. KPS	Informasi mengenai nomor kartu perlindungan sosial
25	Nama Ayah	Informasi mengenai nama ayah dari calon santri
26	NIK Ayah	Informasi mengenai nomor induk kependudukan dari ayah dari calon santri
27	Pekerjaan Ayah	Informasi mengenai pekerjaan ayah dari calon santri
28	Penghasilan Ayah	Informasi mengenai penghasilan ayah dari calon santri
29	Nama Ibu	Informasi mengenai nama ibu dari calon santri
30	NIK Ibu	Informasi mengenai nomor induk kependudukan dari ibu dari calon santri
31	Pekerjaan Ibu	Informasi mengenai pekerjaan ibu dari calon santri
32	Penghasilan Ibu	Informasi mengenai penghasilan ibu dari calon santri
33	Nama Wali	Informasi mengenai nama wali dari calon santri
33	NIK Wali	Informasi mengenai nomor induk kependudukan dari wali dari calon santri
34	Pekerjaan Wali	Informasi mengenai pekerjaan wali dari calon santri

35	Penghasilan Wali	Informasi mengenai penghasilan wali dari calon santri
36	No.Hp	Informasi mengenai nomor handphone calon santri
37	Email	Informasi mengenai email calon santri
38	Tinggi	Informasi mengenai tinggi badan calon santri
39	Berat	Informasi mengenai berat badan calon santri

3.2 Data Preprocessing

Pada tahapan ini, akan dilakukan preprocessing dataset dengan tujuan agar menghasilkan kualitas dataset dan analisis yang baik. Adapun preprocessing dilakukan dengan beberapa tahap. Adapun tahapan-tahapan yang akan dilakukan sebagai berikut:

2.1.1 Data Reduction

Tahapan reduction merupakan tahapan untuk menghapus variabel yang tidak akan digunakan dalam analisis karena variabel tersebut tidak selaras dengan arah tujuan analisis. Penentuan variable tersebut berdasarkan wawancara yang telah dilakukan. adapun variabel-variabel yang digunakan dalam analisis seperti pada gambar 2 dibawah ini:

Tabel 2 Hasil Reduction Data

Hasil <i>Reduction</i> Data THN AJARAN, GELOMBANG, JENJANG, JENIS DAFTAR, JENIS KELAMIN, KAB/KOTA, PROVINSI, PEKERJAAN AYAH, PENGHASILAN AYAH, PEKERJAAN IBU, PENGHASILAN IBU, PEKERJAAN WALI, PENGHASILAN WALI, TINGGI, BERAT, ASAL SEKOLAH DAN EKSKUL PILIHAN

2.1.2 Data Transformation

Dalam dataset yang akan dianalisis, terdapat beberapa variabel yang sulit untuk dipahami, hal tersebut bisa menyebabkan gagalnya interpretasi data atau kurang efektifnya insight yang akan didapatkan. Maka daripada itu diperlukannya transformation data untuk memudahkan intrepretasi data dan membuat insight yang tepat. adapun variabel-variabel yang di transformasikan seperti pada tabel 3:

Tabel 3 *Reduction* Data

No	Sebelum Tranformasi Variabel	Sesudah Tranformasi Variabel
1	TH.AJARAN	THN AJARAN
2	GEL	GELOMBANG
3	JK	JENIS KELAMIN
4	KAB./KOTA	KAB/KOTA
5	PROPINSI	PROVINSI

2.1.3 Data Cleaning

Dalam dataset yang diproses terdapat beberapa data yang memiliki karakter tidak diperlukan [^a-zA-Z0-9]. Hal tersebut dapat menyebabkan ketidak efektifan dalam proses analisis data. Maka daripada itu dataset yang memiliki karakter yang tidak diperlukan perlu di

cleaning agar menghasilkan data yang baik. Adapun variabel yang memiliki karakter yang tidak diperlukan sebagai berikut:

Tabel 4 Jumlah *Data Cleaning*

No	Nama Variabel	Jumlah yang di Cleaning
1	JENIS DAFTAR	4308
2	KAB/KOTA	974
3	PROVINSI	3303
4	TINGGI	4
5	BERAT	1
6	ASAL SEKOLAH	3053

2.1.4 Removing Duplicates

Dalam dataset pendaftaran, terdeteksi data duplikat didalamnya. Hal tersebut bisa menyebabkan tidak efektifnya analisi yang dilakukan. Maka diperlukan tindakan untuk mengatasi duplikat dengan cara menghapus data tersebut. Adapun hasil indikasi dalam pencarian di dataset seperti pada gambar 2 dibawah ini:

Jumlah baris duplikat : 284

Gambar 2 Jumlah Data Duplikat

3.3 Exploratory Data Analysis

Exploratory data analysis dilakukan untuk menganalisis dan menginterpretasi data pendaftaran. Didalamnya terdapat tahapan tahapan yang akan dilakukan untuk menganalisis dan interpretasi data. Adapun tahapan yang dilakukan, sebagai berikut:

3.3.1 Central Tendency

Nilai-nilai *central tendency* seperti *mean*, *median*, *modus*, dan *skewness* memiliki peran penting, terutama dalam menangani missing value. Dalam dataset pendaftaran terdapat dua jenis kolom yaitu kolom numerik dan kolom kategorikal untuk kolom numerik *central tendency* nya yaitu mean, median, modus dan skweness dan untuk kolom kategorikal *central tendency*nya modus. Adapun *central tendency* pada dataset sebagai berikut:

Tabel 5 *Central Tendency*

Jenis Kolom	Kolom Data	Jenjang	Mean	Median	Modus	Skewness
Kolom Numerik	TINGGI	SMAT	153.690	150	150	1.11
		SMPT	144.351	140	140	0.29
	BERAT	SMAT	41.032	33	31	1.07
		SMPT	38.509	34	34	1.64
Kolom Kategorikal	PEKERJAN AYAH	SMAT	-	-	Wiraswasta	-
		SMPT	-	-	Wiraswasta	-
	PENGHASILAN AYAH	SMAT	-	-	2.000.000-5.000.000	-
		SMPT	-	-	2.000.000-5.000.000	-
	PEKERJA	SMAT	-	-	Ibu Rumah	-

	AN IBU				Tangga	
		SMPT	-	-	Ibu Rumah Tangga	-

3.3.2 Mengatasi *Missing Value*

Dalam dataset pendaftaran terdapat missing value. Adanya missing value dalam dataset pendaftaran disebabkan oleh beberapa faktor misal karena *human error* dan *system error*. Dalam menangani *missing value* jenis data numerik dan kategorikal menggunakan *central tendency* yang telah dibuat dan menggunakan imputasi nilai khusus. *Skewness*, *mean*, dan *median* digunakan untuk mengisi *missing value* pada data numerik. Dan mengatasi missing value pada kolom kategorikal menggunakan imputasi nilai *modus* atau imputasi dengan nilai khusus dengan disesuaikan dari definisi kolom atau pemahaman kolom. Adapun missing value yang diatasi dalam dataset pendaftaran, seperti pada tabel 6 dibawah ini:

Tabel 6 Mengatasi Missing Value

Kolom yang terdapat Missing Value	Jumlah Missing Value	Jenis Kolom	Mengatasi Missing Value
PEKERJAAN AYAH	1224	Kategorikal	Menggunakan Nilai Modus
PENGHASILAN AYAH	1743	Kategorikal	Menggunakan Nilai Modus
PEKERJAAN IBU	1298	Kategorikal	Menggunakan Imputasi Nilai
PENGHASILAN IBU	2159	Kategorikal	Menggunakan Imputasi Nilai
EKSKUL PILIHAN	2435	Kategorikal	Menggunakan Imputasi Nilai
TINGGI	1200	Numerik	Menggunakan Nilai Median
BERAT	1098	Numerik	Menggunakan Nilai Median

3.3.3 Mengatasi *Outlier*

Outlier merupakan nilai-nilai yang berbeda secara signifikan dari sebagian besar data dalam sebuah dataset. Dalam dataset pendaftaran, keberadaan outlier dapat menunjukkan adanya anomali atau kesalahan pengukuran cara mencari outlier dengan menentukan range minimum dan maximum pada suatu kolom. indikator data outliers pada dataset pendaftaran berdasarkan kolom tinggi, berat dan jenjang. Kolom jenjang untuk membagi kategori outlier menjadi data outlier SMAT dan SMPT dan kolom tinggi dan berat untuk menjadi indikator data yang outlier dari definisi kolom tersebut. Adapun outlier yang terindikasi pada dataset pendaftaran seperti pada tabel 7 dibawah ini:

Tabel 7 Outlier

Nama kolom	Range Minimum	Range Maximum	Jumlah Outlier
Tinggi SMPT	<90	>190	20
Tinggi SMAT	<100	>200	6
Berat SMPT	<15	>95	16
Berat SMAT	<25	>100	8

Tabel 7 diatas menjelaskan jumlah outlier, outlier tersebut di dapatkan berdasarkan range minimum dan maximum. Penentuan range minimum dan maximum tersebut didapatkan berdasarkan penelitian oleh Ditya Yantikusuma S yang menyatakan bahwa range minimum dan maximum tinggi dan berat untuk siswa SMAT dan SMPT[9].

3.3.4 Analisis dan Interpretasi

Dalam dataset pendaftaran, terdapat tiga dimensi untuk menganalisis dan menginterpretasi pada dataset pendaftaran. tiga dimensi pada dataset pendaftaran yaitu *univariate*, *bivariate* dan *multivariate*. Adapun untuk penjelasannya, sebagai berikut:

1) Univariate

Dalam melakukan analisis dan interpretation di dataset pendaftaran dilakukannya dimensi univariate yaitu proses analisis dan interpretasi yang fokus pada satu kolom atau variabel. Adapun analisis dan interpretasi univariate sebagai berikut:

- *Univariate Thn Ajaran*

Tabel 8 Univariate Thn Ajaran

THN AJARAN	Jumlah	Tahun
2019/2020	824	2019
2020/2021	1027	2020
2021/2022	1419	2021
2022/2023	1073	2022

Tabel 8 memperlihatkan interpretasi univariate untuk kolom "THN AJARAN." Dalam Tabel 4.9, jumlah pendaftar atau santri dalam beberapa tahun ajaran dipresentasikan dalam kategori tahun. Pada tahun ajaran 2019/2020, terdapat 824 pendaftar, mengalami lonjakan signifikan menjadi 1027 santri pada tahun ajaran 2020/2021. Peningkatan berlanjut pada tahun ajaran 2021/2022, mencapai 1419 santri, menunjukkan pertumbuhan yang stabil. Meskipun mengalami penurunan pada tahun ajaran 2022/2023 menjadi 1073 santri, jumlah pendaftar masih tinggi. Secara keseluruhan, terdapat pertumbuhan yang signifikan dari tahun ke tahun, mencapai puncak pada tahun ajaran 2021/2022, menunjukkan minat yang stabil meskipun terdapat penurunan pada tahun terakhir.

- *Univariate Gelombang*

Tabel 9 Univariate Gelombang

GELOMBANG	Jumlah
1	2154
2	1601
3	588

Tabel 9 memperlihatkan persentase pendaftar di setiap gelombang. Dalam gambar tersebut menjelaskan dari semua pendaftar yang ada kebanyakan santri mendaftar di gelombang pertama dengan angka 2154 santri, gelombang pendaftar terbanyak kedua yaitu di gelombang kedua dengan angka 1601 santri, dan yang terakhir di gelombang ketiga dengan angka 588 santri.

- *Univariate Kab/Kota*

Tabel 10 Univariate Kab/Kota

No	KAB/KOTA	Jumlah
1	Tasikmalaya	2341
2	Bandung	413
3	Garut	310
4	Ciamis	218
5	Bekasi	118
.....
107	Lampung Barat	1
108	Lampung Utara	1
109	Lebak	1
110	Madiun	1
111	Klungkung	1

Tabel 11 diatas memperlihatkan asal Kabupaten dan Kota peserta yang mendaftar dengan jumlah 112 kabupaten dan kota asal peserta yang mendaftar. Dari jumlah keseluruhan kabupaten dan kota yang mendaftar terdapat top 5 kabupaten dan kota yang konsisten selalu mendaftar di setiap tahunnya yaitu Tasikmalaya dengan jumlah 2341 orang, Bandung dengan jumlah 413 orang, Garut dengan jumlah 310 orang, Ciamis dengan jumlah 218 orang dan Bekasi dengan jumlah 118 orang.

- *Univariate Provinsi*

Tabel 11 Univariate Provinsi

PROVINSI	Jumlah
Jawa Barat	4192
Dki Jakarta	49
Banten	42
Jawa Tengah	11
Sumatera Selatan	8

Tabel 12 diatas memperlihatkan jumlah asal provinsi peserta pendaftar dengan jumlah 24 provinsi asal peserta yang mendaftar. Dari jumlah keseluruhan provinsi yang mendaftar terdapat top 5 provinsi yang konsisten selalu mendaftar disetiap tahunnya yaitu Jawa Barat dengan jumlah 4192 orang, DKI Jakarta dengan jumlah 49 orang, Banten dengan jumlah 42 orang, Jawa Tengah dengan jumlah 11 orang Dan Sumatera Selatan dengan jumlah 8 orang.

2) *Bivariate*

Dalam melakukan analisis dan interpretation di dataset pendaftaran dilakukannya dimensi bivariate yaitu proses analisis dan interpretasi yang fokus pada dua kolom atau variabel. Adapun analisis dan interpretasi bivariate sebagai berikut:

- *Bivariate Thn Ajaran dam Jenjang*

Tabel 12 Bivariate Thn Ajaran dam Jenjang

THN AJARAN	Jumlah SMAT	Jumlah SMPT
2019/2020	253	571
2020/2021	556	471
2021/2022	662	757
2022/2023	313	760

Tabel 13 diatas memperlihatkan Tren pendaftaran pada kedua jenjang mengalami fluktuasi dari tahun ke tahun. Terdapat variasi yang signifikan antara jumlah pendaftar pada kedua jenjang, dengan lonjakan dan penurunan yang tidak konsisten dari tahun ke tahun. Hal ini dapat dipengaruhi oleh berbagai faktor, seperti kebijakan penerimaan santri, perubahan minat santri, atau citra sekolah pada periode tertentu.

- *Bivariate Gelombang dan Jenjang*

Tabel 13 Bivariate Gelombang dan Jenjang

Gelombang	Jumlah SMAT	Jumlah SMPT
1	835	1319
2	724	877
3	225	363

Tabel 14 memberikan gambaran hubungan antara gelombang ujian masuk dan jumlah santri di SMAT dan SMPT. Pada gelombang pertama, terdapat partisipasi tinggi dengan 835 santri di SMAT dan 1319 santri di SMPT. Gelombang kedua menunjukkan penurunan dengan 724 santri di SMAT dan 877 santri di SMPT. Gelombang ketiga menunjukkan penurunan yang signifikan, hanya ada 225 santri di SMAT dan 363 santri di SMPT. Pola ini menggambarkan variasi dalam partisipasi santri di kedua sekolah, dengan penurunan mencolok pada gelombang ketiga.

- *Bivariate Jenis Kelamin dan Jenjang*

Tabel 14 Bivariate Jenis Kelamin dan Jenjang

JENIS KELAMIN	Jumlah SMAT	Jumlah SMPT
Perempuan	1267	1514
Laki-Laki	517	1045

Tabel 15 memperlihatkan pola yang tergambar menunjukkan bahwa jumlah perempuan yang mendaftar ujian masuk cenderung lebih tinggi daripada jumlah laki-laki pada kedua jenis sekolah, baik SMAT maupun SMPT. Perbandingan antara jumlah perempuan dan laki-laki yang mendaftar di SMPT juga menunjukkan perbedaan yang signifikan, dengan jumlah perempuan yang lebih tinggi secara signifikan dibandingkan jumlah laki-laki.

3) *Multivariate*

Dalam melakukan analisis dan interpretation di dataset pendaftaran dilakukannya dimensi multivariate yaitu proses analisis dan interpretasi yang fokus pada tiga atau lebih kolom atau variabel. Adapun analisis dan interpretasi Multivariate sebagai berikut:

- *Multivariate Thn Ajaran, Jenjang dan Ekskul Pilihan*

Tabel 15 Multivariate Thn Ajaran, Jenjang dan Ekskul Pilihan

Thn Ajaran	Peminat Ekskul SMAT	Peminat Ekskul SMPT
2019/2020	253	571
2020/2021	556	471

2021/2022	662	757
2022/2023	313	760

Tabel 16 diatas merupakan jumlah minat ekstrakurikuler dari tahun ke tahun untuk bagi setiap jenjang. Tidak memiliki minat masih menjadi tren dominan di kedua jenjang dari tahun ke tahun. Peningkatan minat terlihat pada beberapa kategori tertentu, terutama Olahraga/Beladiri, Bahasa, dan Pramuka. Meskipun terdapat variasi, belum memiliki minat tetap menjadi tantangan yang signifikan dalam mendorong partisipasi dalam ekstrakurikuler di kedua jenjang ini.

3.4 Data Validasi

Proses validasi yang teradopsi dalam tahap ini mengandalkan alat metode *Cramers V*. Metode ini menjadi sarana efektif untuk memvalidasi hubungan antar tabel, sehingga output akhirnya baik atau tidaknya untuk dianalisis. Bagi nilai *Cramers V* yang mendekati 1 menandakan baik untuk dianalisis tetapi untuk nilainya yang mendekati angka 0 berarti cukup baik untuk di analisis dan untuk yang nilainya 0 tidak baik untuk dianalisis. Adapun hasil validasi dari kolom dataset pendaftaran sebagai berikut:

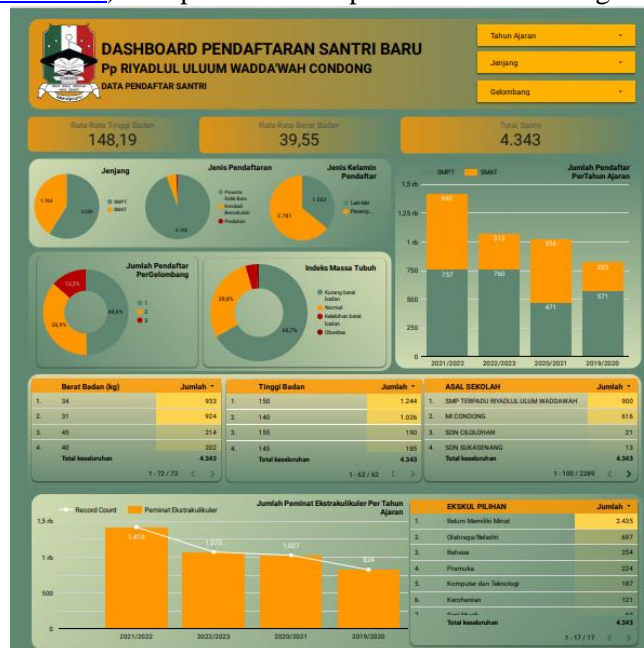
Tabel 16 Validasi Variabel

Variabel yang di Validasi	Nilai Cramers V Sesudah Preprocessing	Keterangan Validasi
THN AJARAN dan JUMLAH	1	Baik Sekali
JENJANG dan JUMLAH	0,999524233	Baik Sekali
JENIS KELAMIN dan JUMLAH	0,999499991	Baik Sekali
KAB/KOTA dan JUMLAH	1	Baik Sekali
PROVINSI dan JUMLAH	1	Baik Sekali
PEKERJAAN AYAH dan JUMLAH	1	Baik Sekali
PENGHASILAN AYAH dan JUMLAH	1	Baik Sekali
PEKERJAAN IBU dan JUMLAH	1	Baik Sekali
PENGHASILAN IBU dan JUMLAH	1	Baik Sekali
PEKERJAAN WALI dan JUMLAH	1	Baik Sekali
PENGHASILAN WALI dan JUMLAH	1	Baik Sekali
ASAL SEKOLAH dan JUMLAH	1	Baik Sekali
EKSKUL PILIHAN dan JUMLAH	1	Baik Sekali
GELOMBANG dan JUMLAH	1	Baik Sekali
TINGGI dan JUMLAH	1	Baik Sekali
BERAT dan JUMLAH	1	Baik Sekali
THN AJARAN dan JENJANG	0,207576978	Baik
GELOMBANG dan JENJANG	0,06077941	Cukup Baik
JENIS KELAMIN dan JENJANG	0,120118713	Baik
KAB/KOTA dan JENJANG	0,056535553	Cukup Baik
PROVINSI dan JENJANG	0,036291749	Cukup Baik
ASAL SEKOLAH dan JENJANG	0,686714624	Baik Sekali
PEKERJAAN AYAH dan JENJANG	0,104095816	Baik
PEKERJAAN IBU dan JENJANG	0,1288671	Baik
PEKERJAAN WALI dan JENJANG	0,129459503	Baik
PENGHASILAN AYAH dan THN AJARAN	0,049761169	Cukup Baik

PENGHASILAN IBU dan THN AJARAN	0,078300404	Cukup Baik
PENGHASILAN WALI dan THN AJARAN	0,04054858	Cukup Baik
TINGGI dan BERAT	0.677500	Baik Sekali
BMI dan JENJANG	0,455068816	Baik
EKSKUL PILIHAN dan JENJANG	0,138695589	Baik
THN AJARAN, JENJANG dan EKSKUL PILIHAN	0.205537	Baik

3.5 Visualisasi Data

Setelah melakukan tahapan *preprocessing* dan *exploratory data analisis*, selanjutnya yaitu tahapan visualisasi data, tahapan tersebut dilakukan agar memudahkan *stakeholder* dalam memahami hasil analisis dan interpretasi. Pada pembuatan dashboard ini dibuat tiga jenis yaitu dashboard khusus data peserta didik, data orangtua dan data sebaran pendaftar (bit.ly/DASHBOARDPSB). Adapun untuk tampilan dashboard sebagai berikut :



Gambar 3 Tampilan Dashboard System Pendaftaran Santri Baru

Agar dashboard dapat dioptimalkan dalam penggunaannya, disini penulis melakukan mini riset untuk mengetahui efektifitas dashboard Penerimaan Santri Baru (PSB). Dari 5 responden yang terdiri dari ketua, sekretaris dan anggota mengatakan efektif untuk dijadikan acuan dalam membantu membangun strategi kegiatan.

3.6 Insight dan Recommendation

Setelah melalui proses analisis, interpretasi, dan memahami pola dashboard, Maka ditemukan beragam *insight* dan rekomendasi yang dapat memperkaya pengambilan keputusan serta perencanaan strategis untuk optimalisasi kegiatan penerimaan santri baru. Temuan tersebut diuraikan dalam tabel berikut ini:

Tabel 17 *Insight dan Recommendation*

Variabel	<i>Insight</i>	Recomendasi
THN AJARAN dan JUMLAH	1) Tren Pertumbuhan 2) Minat Peningkatan	1) Ekspansi Penerimaan. Melakukan strategi berdasarkan data. 2) Penguatan Promosi
JENJANG dan JUMLAH	Preferensi Jenjang Pendidikan	Promosi untuk SMAT ditingkatkan.
JENIS KELAMIN dan JUMLAH	santri perempuan (P) jauh lebih tinggi daripada jumlah santri laki-laki (L).	Strategi Penerimaan yang Lebih Inklusif
KAB/KOTA, JENJANG dan JUMLAH	Pemusatan Pendaftar	Diversifikasi Penerimaan dengan promosi ads
PROVINSI, JENJANG dan JUMLAH	Konsentrasi Pendaftar	Diversifikasi Jangkauan dengan promosi ads
PEKERJAAN AYAH, PEKERJAAN WALI, JENJANG dan JUMLAH	Mayoritas ayah dan wali santri bekerja di sektor swasta (Karyawan Swasta, Wiraswasta)	Kerjasama dengan Komunitas
PENGHASILAN AYAH, PENGHASILAN WALI, JENJANG dan JUMLAH	Mayoritas ayah dan wali santri memiliki penghasilan dalam rentang 2.000.000 hingga 5.000.000	Program Pengembangan Ekonomi bagi yang berpenghasilan rendah
PEKERJAAN IBU dan JUMLAH	Dominasi Ibu Rumah Tangga	Program Edukasi bagi Ibu Rumah Tangga
PENGHASILAN IBU dan JUMLAH	Mayoritas ibu memiliki penghasilan yang berada di kisaran menengah ke bawah	Program Pendidikan Keuangan untuk ibu
ASAL SEKOLAH, JENJANG dan JUMLAH	Asal calon santri baru, dengan sejumlah besar santri berasal dari SMP Terpadu Riyadlul Ulum	Kerjasama Strategis
EKSKUL PILIHAN, JENJANG dan JUMLAH	1) Minat Olahraga/Beladiri yang Tinggi 2) Trend Peminat Ekskul Menunjukkan Fluktuasi	1) Promosi Ekskul yang Lebih Efektif 2) Evaluasi dan Penyesuaian
BMI dan JENJANG	Kategori BMI Kurang Berat Badan yang Tinggi pada SMPT	1) Program Kesehatan dan Nutrisi 2) Pendidikan Gizi

4. KESIMPULAN

Berdasarkan hasil implementasi Data Analytics dalam mengelola dataset pendaftaran dimulai dengan data collecting dan understanding, data preprocessing, exploratory data analysis, validasi data dan visualisasi data menggunakan lookerstudio, maka dapat disimpulkan bahwa:

1. Menghasilkan Dashboard yang memberikan panduan berbasis data bagi stakeholder dalam pengambilan keputusan dan mendukung kebijakan serta strategi yang terukur.
2. Insight dari Dashboard pendaftaran santri baru memberikan rekomendasi untuk optimalisasi kegiatan penerimaan santri baru berdasarkan analisis data. Dashboard tersebut juga sudah dilakukan penilaian oleh pihak terkait dan terbukti efektif.

DAFTAR PUSTAKA

- [1] M. Radhi, A. Amalia, D. R. H. Sitompul, S. H. Sinurat, and E. Indra, "Analisis Big Data Dengan Metode Exploratory Data Analysis (Eda) Dan Metode Visualisasi Menggunakan Jupyter Notebook," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 4, no. 2, pp. 23–27, 2022, doi: 10.34012/journalsisteminformasidanilmukomputer.v4i2.2475.
- [2] D. Hartama, "Analisa Visualisasi Data Akademik Menggunakan Tableau Big Data," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 3, no. 3, p. 46, 2018, doi: 10.30645/jurasik.v3i0.65.
- [3] Heri Purnadi, "Pemanfaatan Google Spreadsheet Dan Google Data Studio Sebagai Dashboard Suhu Dan Kelembaban Di Laboratorium," *Insa. Metrol. PPSDK*, vol. 1, no. 1, pp. 28–33, 2021, doi: 10.55101/ppsdk.v1i1.639.
- [4] Rezkia, "Langkah-Langkah Menggunakan Teknik Analisis Data Kualitatif," *Langkah-Langkah Menggunakan Tek. Anal. Data Kualitatif*, 2020, [Online]. Available: <https://dqlab.id/data-analisis-pahami-teknik-pengumpulan-data>
- [5] A. Rijali, "Analisis Data Kualitatif," *Alhadharah J. Ilmu Dakwah*, vol. 17, no. 33, p. 81, 2019, doi: 10.18592/alhadharah.v17i33.2374.
- [6] F. V. P. Samosir, L. P. Mustamu, E. D. Anggara, A. I. Wiyogo, and A. Widjaja, "Exploratory Data Analysis terhadap Kepadatan Penumpang Kereta Rel Listrik," *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 2, pp. 449–467, 2021, doi: 10.28932/jutisi.v7i2.3700.
- [7] A. Darmawan, N. Kustian, and W. Rahayu, "Implementasi Data Mining Menggunakan Model SVM untuk Prediksi Kepuasan Pengunjung Taman Tabebuaya," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 2, no. 3, p. 299, 2018, doi: 10.30998/string.v2i3.2439.
- [8] A. Jariyah, T. Indrabulan, M. I. Syarif, and R. Krisna, "Pemanfaatan Looker Studio untuk Visualisasi Kinerja Program Studi D4 Teknik Multimedia dan Jaringan Pendahuluan," vol. 1, no. 4, pp. 873–882, 2022.
- [9] D. Yankusuma Setiani and R. Indriati, "Pemeriksaan Status Gizi Anak Sekolah Sebagai Upaya Pencegahan Masalah Gizi," *Abdimas Kosala J. Pengabd. Kpd. Masy.*, vol. 2, no. 2, pp. 63–69, 2023, doi: 10.37831/akj.v2i2.286.