

UAS Kapita Selekta Statistika

Kelompok 10 (Butuh Libur) - Kelas A

Dennis Immanuel

6161901066

Putu Lila Kusumayani

6161901104

Bilqis Aulia R.

6161901114

Tiara Alamanda

6161901116

BIG DATA PIPELINE ETL WITH APACHE SPARK SQL ANALYSIS AND VISUALIZATION WITH LOOKER STUDIO FOR 2019 UTBK SCORE



- PENDAHULUAN -

Ujian Tulis Berbasis Komputer (UTBK) merupakan ujian seleksi masuk perguruan tinggi negeri yang paling bergengsi. Tes tersebut diadakan setiap tahun menjelang masa penerimaan mahasiswa baru dan diikuti oleh puluhan ribu peserta didik. Pada penelitian ini, kita akan mengamati profil peserta, universitas, dan jurusan pilihan yang terdaftar pada seleksi UTBK, untuk menganalisis tren dan performa SBMPTN yang diselenggarakan pada tahun 2019.



D A T A

Data di-ekstrak dari salah satu datasets **kaggle** yang berjudul **Indonesia College Entrance Examination - UTBK 2019**. Dengan URL Link : <https://www.kaggle.com/datasets/ekojsalim/indonesia-college-entrance-examination-utbk-2019>

| | id_major | id_university | type | major_name | capacity |
|------|-----------------|----------------------|-------------|-----------------------------|-----------------|
| 0 | 1111014 | 111 | science | PENDIDIKAN DOKTER HEWAN | 88 |
| 1 | 1111022 | 111 | science | TEKNIK SIPIL | 64 |
| 2 | 1111037 | 111 | science | TEKNIK MESIN | 48 |
| 3 | 1111045 | 111 | science | TEKNIK KIMIA | 48 |
| 4 | 1111053 | 111 | science | ARSITEKTUR | 48 |
| ... | ... | ... | ... | ... | ... |
| 3162 | 9212011 | 921 | humanities | EKONOMI PEMBANGUNAN | 32 |
| 3163 | 9212042 | 921 | humanities | MANAJEMEN | 32 |
| 3164 | 9212057 | 921 | humanities | AKUNTANSI | 32 |
| 3165 | 9212065 | 921 | humanities | PENDIDIKAN BAHASA INDONESIA | 16 |
| 3166 | 9212073 | 921 | humanities | PENDIDIKAN BAHASA INGGRIS | 16 |

3167 rows × 5 columns

majors.csv

| | id_university | university_name |
|-----|----------------------|-----------------------------|
| 0 | 111 | UNIVERSITAS SYIAH KUALA |
| 1 | 112 | UNIVERSITAS MALIKUSSALEH |
| 2 | 113 | UNIVERSITAS TEUKU UMAR |
| 3 | 114 | UNIVERSITAS SAMUDRA |
| 4 | 115 | ISBI ACEH |
| ... | ... | ... |
| 80 | 821 | UNIVERSITAS KHAIRUN |
| 81 | 911 | UNIVERSITAS CENDERAWASIH |
| 82 | 912 | UNIVERSITAS MUSAMUS MERAUKE |
| 83 | 913 | ISBI TANAH PAPUA |
| 84 | 921 | UNIVERSITAS PAPUA |

85 rows × 2 columns

universities.csv

DATA

→

| | Unnamed: 0 | id_first_major | id_first_university | id_second_major | id_second_university | id_user | score_eko | score_geo | score_kmb | score_kpu | score_kua | score_lis | score_sci |
|-------|---------------|----------------|---------------------|-----------------|----------------------|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0 | 0 | 3322014 | 332 | 3532154 | | 353 | 26 | 778 | 486 | 679 | 594 | 643 | |
| 1 | 1 | 3212057 | 321 | 3322022 | | 332 | 35 | 579 | 631 | 710 | 700 | 810 | |
| 2 | 2 | 3722057 | 372 | 1912042 | | 191 | 44 | 600 | 481 | 314 | 432 | 522 | |
| 3 | 3 | 3212081 | 321 | 3212301 | | 321 | 55 | 559 | 563 | 646 | 638 | 591 | |
| 4 | 4 | 3812106 | 381 | 3852034 | | 385 | 171 | 838 | 734 | 654 | 613 | 690 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 61197 | 61197 | 3552127 | 355 | 6212115 | | 621 | 328356 | 521 | 497 | 498 | 522 | 573 | |
| 61198 | 61198 | 3512053 | 351 | 3532057 | | 353 | 328357 | 638 | 464 | 548 | 576 | 541 | |
| 61199 | 61199 | 6212115 | 621 | 6212107 | | 621 | 328364 | 491 | 601 | 567 | 486 | 492 | |
| 61200 | 61200 | 3812056 | 381 | 3822012 | | 382 | 328366 | 573 | 544 | 634 | 672 | 527 | |
| 61201 | 61201 | 3562362 | 356 | 3562192 | | 356 | 328373 | 561 | 627 | 468 | 457 | 530 | |

score_humanities.csv

| | Unnamed: 0 | id_first_major | id_first_university | id_second_major | id_second_university | id_user | score_bio | score_fis | score_kim | score_kmb | score_kpu | score_kva |
|-------|---------------|----------------|---------------------|-----------------|----------------------|---------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0 | 0 | 3321065 | 332 | 3331187 | 333 | 4 | 400 | 400 | 400 | 400 | 400 | 400 |
| 1 | 1 | 3211015 | 321 | 3611066 | 361 | 14 | 816 | 666 | 651 | 678 | 685 | |
| 2 | 2 | 3721093 | 372 | 3551302 | 355 | 19 | 562 | 839 | 624 | 700 | 781 | |
| 3 | 3 | 3321096 | 332 | 3551194 | 355 | 23 | 700 | 669 | 692 | 679 | 692 | |
| 4 | 4 | 5211104 | 521 | 5211085 | 521 | 28 | 461 | 619 | 441 | 593 | 563 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 86565 | 86565 | 1411161 | 141 | 1421147 | 142 | 344111 | 516 | 412 | 376 | 546 | 523 | |
| 86566 | 86566 | 3561262 | 356 | 3571055 | 357 | 344125 | 334 | 634 | 435 | 613 | 461 | |
| 86567 | 86567 | 3531182 | 353 | 3561022 | 356 | 344127 | 316 | 717 | 421 | 474 | 494 | |
| 86568 | 86568 | 1121094 | 112 | 1221095 | 122 | 344151 | 592 | 451 | 583 | 437 | 447 | |
| 86569 | 86569 | 1911286 | 191 | 1921056 | 192 | 344192 | 411 | 602 | 415 | 519 | 540 | |



Hasil Analisis dan Pembahasan

ANALYSIS WITH PYSPARK SQL

1 Jurusan Paling diminati dari penjurusan IPA & IPS

```
favorit_ipa = spark.sql("""
    SELECT Nama_Jurusan, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT major_name_x AS Nama_Jurusan FROM data_ipa
        UNION ALL
        SELECT major_name_y AS Nama_Jurusan FROM data_ipa
    ) AS combined_data
    GROUP BY Nama_Jurusan
    ORDER BY Jumlah_Pendaftar DESC
    LIMIT 10;
"""

)
favorit_ipa.limit(favorit_ipa.count()).show(truncate=False)
```

| Nama_Jurusan | Jumlah_Pendaftar |
|--------------------|------------------|
| PENDIDIKAN DOKTER | 9415 |
| TEKNIK SIPIL | 6299 |
| FARMASI | 5793 |
| TEKNIK INFORMATIKA | 4889 |
| TEKNIK INDUSTRI | 4838 |
| BIOLOGI | 4516 |
| AGRIBISNIS | 4242 |
| TEKNIK ELEKTRO | 3976 |
| MATEMATIKA | 3925 |
| KIMIA | 3835 |

Top 10 jurusan paling diminati dari penjurusan IPA

1A

Top 10 jurusan paling diminati dari penjurusan IPS

1B

```
favorit_ips = spark.sql("""
    SELECT Nama_Jurusan, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT major_name_x AS Nama_Jurusan FROM data_ips
        UNION ALL
        SELECT major_name_y AS Nama_Jurusan FROM data_ips
    ) AS combined_data
    GROUP BY Nama_Jurusan
    ORDER BY Jumlah_Pendaftar DESC
    LIMIT 10;
"""

)
favorit_ips.limit(favorit_ips.count()).show(truncate=False)
```

| Nama_Jurusan | Jumlah_Pendaftar |
|---------------------|------------------|
| MANAJEMEN | 10492 |
| ILMU HUKUM | 8345 |
| ILMU KOMUNIKASI | 7938 |
| AKUNTANSI | 7468 |
| PSIKOLOGI | 5628 |
| EKONOMI PEMBANGUNAN | 3603 |
| SOSIOLOGI | 3335 |
| HUKUM | 2797 |
| SASTRA INGGRIS | 2369 |
| ILMU POLITIK | 2131 |

ANALYSIS WITH PYSPARK SQL

2 Top 10 Universitas paling diminati (baik dari penjurusan IPA maupun IPS)

```
favorit_univ1 = spark.sql("""
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT university_name_x AS Nama_Universitas FROM data_ipa
        UNION ALL
        SELECT university_name_y AS Nama_Universitas FROM data_ipa -----)
    ) AS combined_data
    GROUP BY Nama_Universitas
    ORDER BY Jumlah_Pendaftar DESC;
""")

.createOrReplaceTempView("favorit_univ1")
```

i

```
favorit_univ2 = spark.sql("""
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT university_name_x AS Nama_Universitas FROM data_ips
        UNION ALL
        SELECT university_name_y AS Nama_Universitas FROM data_ips -----)
    ) AS combined_data
    GROUP BY Nama_Universitas
    ORDER BY Jumlah_Pendaftar DESC;
""")

.createOrReplaceTempView("favorit_univ2")
```

ii

```
favorit_univ = spark.sql("""
    SELECT Nama_Universitas, SUM(Jumlah_Pendaftar) AS Jumlah_Pendaftar
    FROM (
        SELECT Nama_Universitas, Jumlah_Pendaftar
        FROM (
            SELECT favorit_univ1.Nama_Universitas, favorit_univ1.Jumlah_Pendaftar
            FROM favorit_univ1
            UNION ALL
            SELECT favorit_univ2.Nama_Universitas, favorit_univ2.Jumlah_Pendaftar
            FROM favorit_univ2 -----)
        ) AS combined_data
    ) AS subquery
    GROUP BY Nama_Universitas
    ORDER BY Jumlah_Pendaftar DESC
    LIMIT 10
""")

favorit_univ.limit(favorit_univ.count()).show(truncate=False)
```

iii

| Nama_Universitas | Jumlah_Pendaftar |
|----------------------------------|------------------|
| UNIVERSITAS BRAWIJAYA | 18839 |
| UNIVERSITAS DIPONEGORO | 16193 |
| UNIVERSITAS PADJADJARAN | 15366 |
| UNIVERSITAS GADJAH MADA | 13836 |
| UNIVERSITAS SEBELAS MARET | 13783 |
| UNIVERSITAS INDONESIA | 13716 |
| UNIVERSITAS PENDIDIKAN INDONESIA | 10427 |
| UNIVERSITAS JENDERAL SOEDIRMAN | 9156 |
| UNIVERSITAS NEGERI SEMARANG | 8617 |
| UNIVERSITAS JEMBER | 8416 |

Results

ANALYSIS WITH PYSPARK SQL

3

Berdasarkan jurusan paling diminati universitas mana yang paling banyak peminatnya?

```
favorit_univ_dokter = spark.sql("""
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT university_name_x AS Nama_Universitas FROM data_ipa
        UNION ALL
        SELECT university_name_y AS Nama_Universitas FROM data_ipa
        WHERE data_ipa.major_name_x = "PENDIDIKAN DOKTER" OR data_ipa.major_name_y = "PENDIDIKAN DOKTER"
    ) AS combined_data
    GROUP BY Nama_Universitas
    ORDER BY Jumlah_Pendaftar DESC
    LIMIT 10;
"""
)
favorit_univ_dokter.limit(favorit_univ_dokter.count()).show(truncate=False)
```

| Nama_Universitas | Jumlah_Pendaftar |
|-------------------------------------|------------------|
| UNIVERSITAS BRAWIJAYA | 6611 |
| UNIVERSITAS GADJAH MADA | 5792 |
| UNIVERSITAS DIPONEGORO | 5596 |
| UNIVERSITAS PADJADJARAN | 5114 |
| UNIVERSITAS INDONESIA | 4741 |
| UNIVERSITAS SEBELAS MARET | 4105 |
| INSTITUT PERTANIAN BOGOR | 3982 |
| INSTITUT TEKNOLOGI SEPULUH NOPEMBER | 3435 |
| INSTITUT TEKNOLOGI BANDUNG | 3371 |
| UNIVERSITAS JENDERAL SOEDIRMAN | 3009 |

3A Top 10 Universitas Paling Diminati untuk Pemilihan Jurusan Favorit IPA (PENDIDIKAN DOKTER)

ANALYSIS WITH PYSPARK SQL

3

Berdasarkan jurusan paling diminati universitas mana yang paling banyak peminatnya?

```
favorit_univ_manajemen = spark.sql("""
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT university_name_x AS Nama_Universitas FROM data_ips
        UNION ALL
        SELECT university_name_y AS Nama_Universitas FROM data_ips
        WHERE data_ips.major_name_x = "MANAJEMEN" OR data_ips.major_name_y = "MANAJEMEN"
    ) AS combined_data
    GROUP BY Nama_Universitas
    ORDER BY Jumlah_Pendaftar DESC
    LIMIT 10;
"""
)
favorit_univ_manajemen.limit(favorit_univ_manajemen.count()).show(truncate=False)
```

| Nama_Universitas | Jumlah_Pendaftar |
|----------------------------------|------------------|
| UNIVERSITAS INDONESIA | 4767 |
| UNIVERSITAS PADJADJARAN | 4173 |
| UNIVERSITAS BRAWIJAYA | 3847 |
| UNIVERSITAS PENDIDIKAN INDONESIA | 3763 |
| UNIVERSITAS DIPONEGORO | 3734 |
| UNIVERSITAS GADJAH MADA | 3449 |
| UNIVERSITAS SEBELAS MARET | 3134 |
| UNIVERSITAS NEGERI JAKARTA | 2986 |
| UNIVERSITAS NEGERI YOGYAKARTA | 2665 |
| UNIVERSITAS NEGERI SEMARANG | 2566 |

3B Top 10 Universitas Paling Diminati untuk Pemilihan Jurusan Favorit IPS (MANAJEMEN)

ANALYSIS WITH PYSPARK SQL

4 Informasi masing-masing Universitas

- Jumlah mahasiswa baru yang diterima tahun ini (kapasitas)
- Rata-rata Skor UTBK pendaftar universitas tersebut
- Persentase penerimaan Universitas

4A Informasi masing-masing universitas untuk segmen penjurusan IPA

```
pendaftar_univ_ipa = spark.sql("""
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT university_name_x AS Nama_Universitas FROM data_ipa
        UNION ALL
        SELECT university_name_y AS Nama_Universitas FROM data_ipa
    ) AS combined_data
    GROUP BY Nama_Universitas
    ORDER BY Jumlah_Pendaftar DESC;
"""

).createOrReplaceTempView("pendaftar_univ_ipa")
```

```
pendaftar_univ_ipa = spark.sql("""
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar
    FROM (
        SELECT university_name_x AS Nama_Universitas FROM data_ipa
        UNION ALL
        SELECT university_name_y AS Nama_Universitas FROM data_ipa
    ) AS combined_data
    GROUP BY Nama_Universitas
    ORDER BY Jumlah_Pendaftar DESC;
"""

).createOrReplaceGlobalTempView("pendaftar_univ_ipa")
```

```
rata2_skor_univ_ipa = spark.sql("""
    SELECT Nama_Universitas, CAST(AVG(rata2_skor_utbk) AS INT) AS skor_utbk
    FROM (
        SELECT university_name_x AS Nama_Universitas, score_utbk AS rata2_skor_utbk FROM data_ipa
        UNION ALL
        SELECT university_name_y AS Nama_Universitas, score_utbk AS rata2_skor_utbk FROM data_ipa
    ) AS combined_data
    GROUP BY Nama_Universitas
    ORDER BY skor_utbk DESC;
"""

).createOrReplaceTempView("rata2_skor_univ_ipa")
```

```
test_kapasitas_univ_ipa = spark.sql("""
    SELECT Nama_Universitas, jurusan, kapasitas
    FROM (
        SELECT university_name_x AS Nama_Universitas, major_name_x AS jurusan, capacity_x AS kapasitas FROM data_ipa
        UNION ALL
        SELECT university_name_y AS Nama_Universitas, major_name_y AS jurusan, capacity_y AS kapasitas FROM data_ipa
    ) AS combined_data
    GROUP BY Nama_Universitas, jurusan, kapasitas
    ORDER BY SUM(kapasitas) DESC;
"""

).createOrReplaceTempView("test_kapasitas_univ_ipa")
```

i

ii

ANALYSIS WITH PYSPARK SQL

4A

Informasi masing-masing universitas untuk segmen penjurusan IPA

```
kapasitas_univ_ipa = spark.sql("""
    SELECT Nama_Universitas, SUM(kapasitas) AS kapasitas
    FROM test_kapasitas_univ_ipa
    GROUP BY Nama_Universitas
    ORDER BY kapasitas DESC;
""")

.createOrReplaceTempView("kapasitas_univ_ipa")

test_info_univ_ipa = spark.sql("""
    SELECT pendaftar_univ_ipa.Nama_Universitas, kapasitas_univ_ipa.kapasitas AS Jumlah_Mahasiswa_Baru, rata2_skor_univ_ipa.skor_utbk AS avg_skor_utbk
    FROM pendaftar_univ_ipa
    JOIN kapasitas_univ_ipa ON pendaftar_univ_ipa.Nama_Universitas = kapasitas_univ_ipa.Nama_Universitas
    JOIN rata2_skor_univ_ipa ON kapasitas_univ_ipa.Nama_Universitas = rata2_skor_univ_ipa.Nama_Universitas
    ORDER BY Jumlah_Mahasiswa_Baru DESC;
""")

.createOrReplaceGlobalTempView("test_info_univ_ipa")

pendaftar_univ_ipa = spark.sql("SELECT * FROM global_temp.pendaftar_univ_ipa")
info_univ_ipa = spark.sql("SELECT * FROM global_temp.test_info_univ_ipa")
```

iii

RUN SEKALI AJA



```
▶ info_univ_ipa = info_univ_ipa.join(pendaftar_univ_ipa, "Nama_Universitas")
info_univ_ipa = info_univ_ipa.withColumn("acceptance_rate", (info_univ_ipa.Jumlah_Mahasiswa_Baru / info_univ_ipa.Jumlah_Pendaftar * 100).cast("double"))
info_univ_ipa = info_univ_ipa.drop("Jumlah_Pendaftar")
info_univ_ipa = info_univ_ipa.orderBy(info_univ_ipa.Jumlah_Mahasiswa_Baru.desc())
info_univ_ipa.limit(info_univ_ipa.count()).show(truncate=False)
```

iv

ANALYSIS WITH PYSPARK SQL

4A

Informasi masing-masing universitas untuk segmen penjurusan IPA

| Nama_Universitas | Jumlah_Mahasiswa_Baru | avg_skor_utbk | acceptance_rate |
|-------------------------------------|-----------------------|---------------|--------------------|
| UNIVERSITAS BRAWIJAYA | 5116 | 557 | 41.83498241884046 |
| UNIVERSITAS DIPONEGORO | 3623 | 569 | 35.48829464198256 |
| UNIVERSITAS GADJAH MADA | 3249 | 587 | 36.25306851149297 |
| UNIVERSITAS JEMBER | 2766 | 540 | 49.810913019989194 |
| UNIVERSITAS PADJADJARAN | 2661 | 559 | 31.44646655636965 |
| UNIVERSITAS SRIWIJAYA | 2587 | 533 | 62.74557361144798 |
| UNIVERSITAS SEBELAS MARET | 2552 | 559 | 31.37062077443147 |
| UNIVERSITAS HASANUDDIN | 2522 | 525 | 69.84214898919967 |
| INSTITUT PERTANIAN BOGOR | 2328 | 558 | 33.24289590175639 |
| UNIVERSITAS AIRLANGGA | 2261 | 565 | 54.972039873571596 |
| UNIVERSITAS UDAYANA | 2220 | 548 | 61.529933481153 |
| UNIVERSITAS TADULAKO | 2127 | 516 | 543.9897698209719 |
| UNIVERSITAS SUMATERA UTARA | 2105 | 533 | 53.69897959183674 |
| INSTITUT TEKNOLOGI SUMATERA | 2051 | 514 | 89.4461404273877 |
| INSTITUT TEKNOLOGI SEPULUH NOPEMBER | 1972 | 584 | 33.209834961266424 |
| UNIVERSITAS INDONESIA | 1832 | 586 | 26.171428571428574 |
| UNIVERSITAS SYIAH KUALA | 1794 | 527 | 91.43730886850153 |
| UNIVERSITAS JENDERAL SOEDIRMAN | 1780 | 543 | 30.47423386406437 |
| UNIVERSITAS ANDALAS | 1764 | 542 | 49.10913140311804 |
| UNIVERSITAS NEGERI MALANG | 1655 | 535 | 52.257657088727505 |

only showing top 20 rows

Results

ANALYSIS WITH PYSPARK SQL

4 Informasi masing-masing Universitas

- Jumlah mahasiswa baru yang diterima tahun ini (kapasitas)
- Rata-rata Skor UTBK pendaftar universitas tersebut
- Persentase penerimaan Universitas

4B Informasi masing-masing universitas untuk segmen penjurusan IPS

```
pendaftar_univ_ips = spark.sql("""  
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar  
    FROM (  
        SELECT university_name_x AS Nama_Universitas FROM data_ips  
        UNION ALL  
        SELECT university_name_y AS Nama_Universitas FROM data_ips  
    ) AS combined_data  
    GROUP BY Nama_Universitas  
    ORDER BY Jumlah_Pendaftar DESC;  
    """  
).createOrReplaceTempView("pendaftar_univ_ips")
```

```
pendaftar_univ_ips = spark.sql("""  
    SELECT Nama_Universitas, COUNT(*) AS Jumlah_Pendaftar  
    FROM (  
        SELECT university_name_x AS Nama_Universitas FROM data_ips  
        UNION ALL  
        SELECT university_name_y AS Nama_Universitas FROM data_ips  
    ) AS combined_data  
    GROUP BY Nama_Universitas  
    ORDER BY Jumlah_Pendaftar DESC;  
    """  
).createOrReplaceGlobalTempView("pendaftar_univ_ips")
```

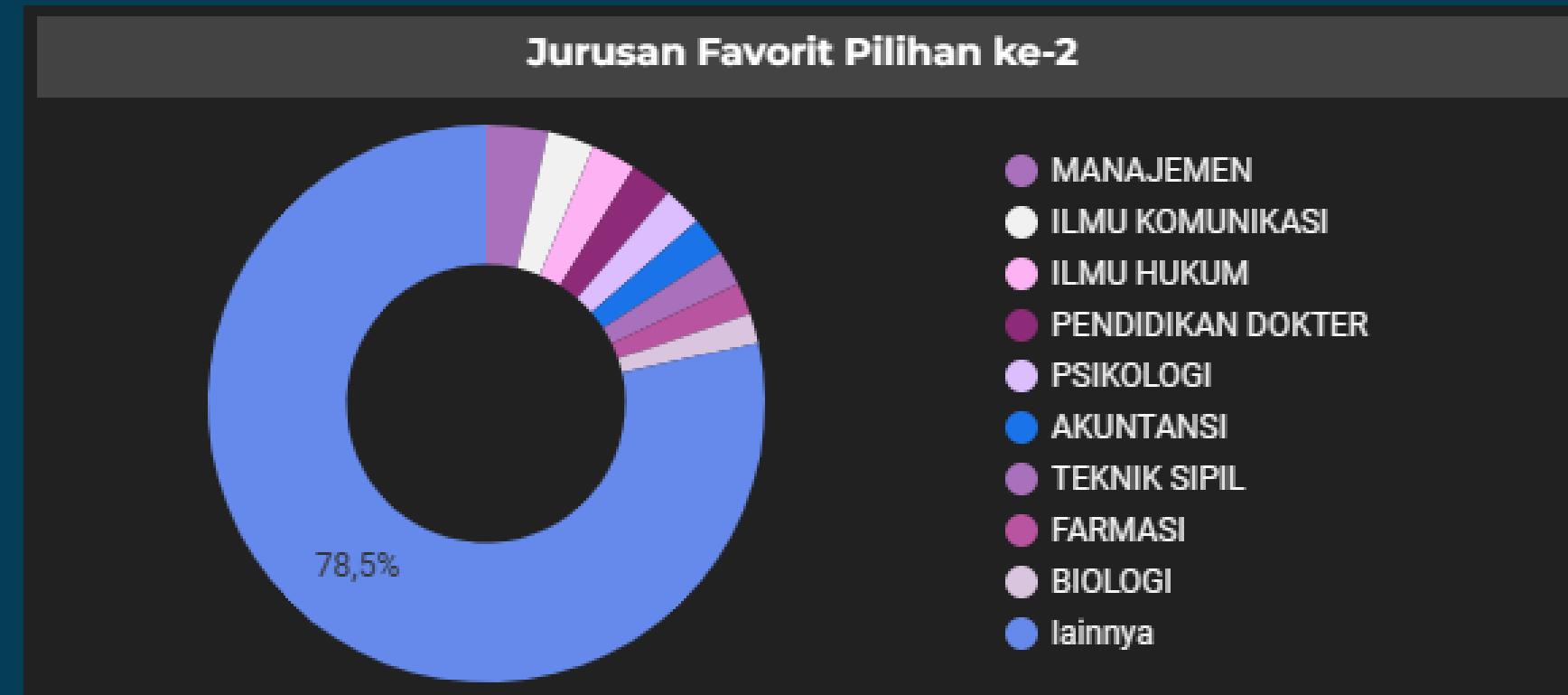
```
rata2_skor_univ_ips = spark.sql("""  
    SELECT Nama_Universitas, CAST(AVG(rata2_skor_utbk) AS INT) AS skor_utbk  
    FROM (  
        SELECT university_name_x AS Nama_Universitas, score_utbk AS rata2_skor_utbk FROM data_ips  
        UNION ALL  
        SELECT university_name_y AS Nama_Universitas, score_utbk AS rata2_skor_utbk FROM data_ips  
    ) AS combined_data  
    GROUP BY Nama_Universitas  
    ORDER BY skor_utbk DESC;  
    """  
).createOrReplaceTempView("rata2_skor_univ_ips")
```

```
test_kapasitas_univ_ips = spark.sql("""  
    SELECT Nama_Universitas, jurusan, kapasitas  
    FROM (  
        SELECT university_name_x AS Nama_Universitas, major_name_x AS jurusan, capacity_x AS kapasitas FROM data_ips  
        UNION ALL  
        SELECT university_name_y AS Nama_Universitas, major_name_y AS jurusan, capacity_y AS kapasitas FROM data_ips  
    ) AS combined_data  
    GROUP BY Nama_Universitas, jurusan, kapasitas  
    ORDER BY SUM(kapasitas) DESC;  
    """  
).createOrReplaceTempView("test_kapasitas_univ_ips")
```

i

ii

Visualisasi dengan Looker

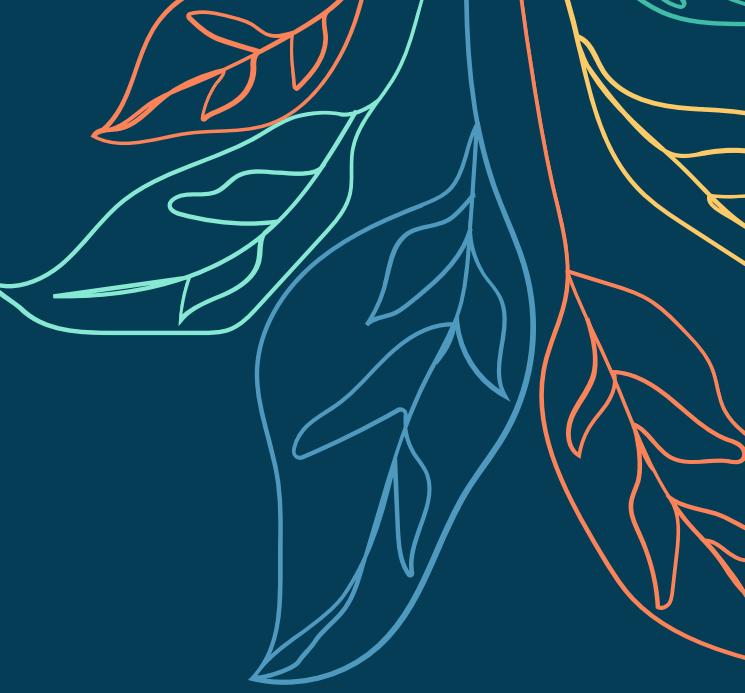


| Jurusan Favorit Pilihan ke-1 | | | | | | | | | | | | | | | | | | |
|------------------------------|----------------|-----------|------------|-----------------|---------------|--------------|-----------|-------------------|--------------|-----------------|---------|-------------|-----------------|------------|---------|--------------|---------|--|
| Semua | | | | | | | | | | | | | | | | | | |
| PENDIDIKAN DO... | | MANAJEMEN | | ILMU HUKUM | | TEKNIK... | | TEKNIK INFORMA... | | SISTEM INFOR... | | TEKNIK E... | | MATEMATIKA | | KESEHATAN... | | |
| PEN... | ILMU... | AKUNTANSI | | ILMU KOMUNIKASI | | EKONOMI P... | | TEKNIK KI... | | STATIST... | | SOSIOL... | | PEND... | | GIZI | | |
| PSIKOLOGI | TEKNIK SIPIL | | FARMASI | | KIMIA ARSI... | | HUKUM | | TEKNIK ME... | | FISIKA | | TEKNIK LINGK... | | ILMU... | | SAST... | |
| PETERN... | AGROTEK... | | INFORMA... | | ILMU PO... | | TE... | | P... | | BI... | | K... | | PE... | | PE... | |
| TEKNOL... | ILMU... | | HUBUNG... | | PENDIDIK... | | TEKNIK... | | DE... | | ILMU... | | PE... | | PEN... | | TE... | |
| TEK... SA... | ILMU ADMINI... | | KEDOKT... | | ILMU D... | | K... | | A... | | FAK... | | BIO... | | BI... | | ILM... | |
| KE... | PENDI... | | TEK... | | PE... | | PEN... | | DE... | | ILMU... | | PE... | | SEK... | | TEKN... | |
| ANTR... | F... | | SE... | | KE... | | P... | | F... | | K... | | P... | | AK... | | T... | |
| | AD... | | ILM... | | I... | | MA... | | E... | | D... | | TE... | | M... | | T... | |
| | MA... | | ILM... | | I... | | MA... | | K... | | S... | | T... | | P... | | P... | |
| | K... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | | I... | |
| | I... | | I...</ | | | | | | | | | | | | | | | |

Visualisasi dengan Looker

| Record Count | | Rata2 mat | Rata2 kmb | Rata2 kpu | Rata2 kua | Rata2 ppu |
|--------------|-----------|-----------|------------|---------------|---------------|---------------|
| 1. | 147767 | 530.15 | 560.09 | 557.41 | 549.15 | 538.8 |
| | | | | | 1 - 1 / 1 | < > |
| Record Count | Rata2 fis | Rata2 kim | Rata2 bio | Rata2 mat IPA | Rata2 kmb... | Rata2 kpu ... |
| 1. | 86569 | 539.34 | 544.6 | 537.14 | 529.49 | 565.8 |
| | | | | | | 1 - 1 / 1 |
| Record... | Rata2 eko | Rata2 geo | Rata2 seja | Rata2 sos | Rata2 mat IPS | Rata2 kmb ... |
| 1. | 61198 | 547.78 | 546.77 | 540.41 | 542.48 | 531.08 |
| | | | | | | 1 - 1 / 1 |

| Perbandingan Pendaftar vs Kapasitas Jurusan | | | |
|---|-----------------------|----------------|------------|
| university_name_x | major_name_x | Recor... | capacit... |
| 1. UNIVERSITAS DIP... | HUKUM | 722 | 424,63 |
| 2. UNIVERSITAS IND... | PENDIDIKAN DOKTER | 596 | 72 |
| 3. INSTITUT TEKNOL... | SEKOLAH TEK. ELEKT... | 550 | 238 |
| 4. UNIVERSITAS IND... | ILMU HUKUM | 543 | 167,08 |
| 5. UNIVERSITAS PAD... | PSIKOLOGI | 531 | 79,96 |
| 6. UNIVERSITAS PAD... | PENDIDIKAN DOKTER | 464 | 124,72 |
| 7. UNIVERSITAS BRA... | ILMU HUKUM | 459 | 259,42 |
| 8. UNIVERSITAS SEB... | KEDOKTERAN | 457 | 109,89 |
| 9. UNIVERSITAS SIND... | VETERINER | 401 | 65 |
| | | 1 - 100 / 2916 | < > |
| university_name_y | major_name_y | Recor... | capacit... |
| 1. UNIVERSITAS DIP... | HUKUM | 473 | 424,22 |
| 2. UPN "VETERAN" J... | ILMU KOMUNIKASI | 385 | 159,79 |
| 3. UNIVERSITAS BRA... | ILMU HUKUM | 371 | 259,58 |
| 4. UNIVERSITAS SEB... | ILMU HUKUM | 354 | 225 |
| 5. UNIVERSITAS BRA... | AGROEKOTEKNOLOGI | 354 | 312 |
| 6. UNIVERSITAS SEB... | KEDOKTERAN | 339 | 109,91 |
| 7. UNIVERSITAS SEB... | PSIKOLOGI | 307 | 60 |
| 8. UPN "VETERAN" .I | MANAJEMEN | 306 | 140,07 |
| | | 1 - 100 / 3101 | < > |



Dengan melakukan analisis seleksi masuk perguruan tinggi melalui nilai UTBK 2019, diharapkan kedepannya analisis ini menjadi salah satu acuan ataupun pertimbangan bagi para mahasiswa dalam menentukan pilihan jurusan dan universitas yang diinginkan agar peluang lulus saat tahap seleksi lebih besar.



#Saran



Untuk proyek kedepannya kelompok kami menyarankan agar saat melakukan visualisasi hasil analisis dapat dilakukan pada *software* lainnya seperti Tableau. Karena pada saat melakukan visualisasi melalui Looker Studio *worksheet* tidak dapat digunakan secara bersamaan dan dapat menyebabkan *error* pada *worksheet* lainnya. Selain itu, Looker Studio juga memerlukan waktu yang cukup lama dalam melakukan pemuatan (*load*) dataset.

-TERIMA
KASIH-