**KLUSTERISASI KUALITAS SUMBER DAYA MANUSIA**

**(Menggunakan K-Means Clustering)**

Mentor:

Mahaning Indrwati Wijaya S. MB., M.Com. CDAP, ITIL

Logo, company name

Description automatically generated

Disusun oleh:

1. Fadhilah Nurbaiti 3373857 / Universitas Islam “45” Bekasi
2. Nindita Ayu Astuti 3799492 / Universitas Slamet Riyadi Surakarta
3. Noven Ramadhani 3407013 / Universitas UPN Veteran Yogyakarta

**Data Analyst & Artificial Intelligence**

**Program Studi Independen Angkatan 3**

**PT. MariBelajar Indonesia Cerdas**

**Tahun 2022**

DAFTAR ISI

[LEMBAR PENGESAHAN 3](#_Toc122035104)

[A. LATAR BELAKANG 4](#_Toc122035105)

[B. PERUMUSAN MASALAH 6](#_Toc122035106)

[C. SPESIFIKASI KEBUTUHAN 6](#_Toc122035107)

[D. RANCANGAN SOLUSI 7](#_Toc122035108)

[E. HASIL DAN PEMBAHASAN 7](#_Toc122035109)

[F. KESIMPULAN 12](#_Toc122035110)

[G. LAMPIRAN 12](#_Toc122035111)

LEMBAR PENGESAHAN

**KLUSTERISASI KUALITAS SUMBER DAYA MANUSIA**

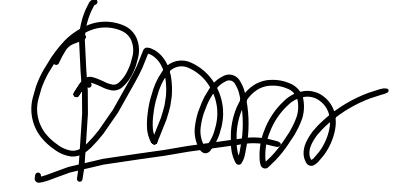
**(Menggunakan K-Means Clustering)**

Disusun oleh:

1. Fadhilah Nurbaiti 3373857 / Universitas Islam “45” Bekasi
2. Nindita Ayu Astuti 3799492 / Universitas Slamet Riyadi Surakarta
3. Noven Ramadhani 3407013 / Universitas UPN Veteran Yogyakarta

Disetujui oleh:

**Mentor**



Mahaning Indrwati Wijaya S. MB., M.Com. CDAP, ITIL

# LATAR BELAKANG

Kualitas sumber daya manusia merupakan komponen penting dalam setiap gerak pembangunan. Hanya dari sumber manusia yang berkualitas tinggi lah yang dapat mempercepat pembangunan bangsa. Jumlah penduduk yang besar, apabila tidak diikuti dengan kualitas yang memadai, hanyalah akan menjadi beban pembangunan. Perlu diketahui bahwa, Kualitas sumber daya manusia itu tidak hanya ditentukan oleh aspek keterampilan atau kekuatan tenaga fisiknya saja, akan tetapi ditentukan oleh Pendidikan, kadar pengetahuannya, bagaimana ia mengelola keuangannya untuk bertahan hidup, pengalaman atau kematangannya dan sikapnya serta nilai-nilai yang dimilikinya. Maka, untuk membantu pemerataan kualitas Sumber Daya Manusia dibuatlah data klusterisasi dari Kualitas SDM untuk mempermudah pemerataan.

Salah satu indikator keberhasilan kualitas sumber daya manusia di suatu wilayah dapat dilihat dari tingkat kualitas sumber daya manusia suatu masyarakat yang dicerminkan melalui Indeks Pembangunan Manusia (IPM). Indeks Pembangunan Manusia (IPM) merupakan pengukuran perbandingan dari harapan hidup, pendidikan, dan standar hidup untuk semua negara. IPM juga menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya [2]. Indeks Pembangunan Manusia digunakan sebagai indikator tolak ukur maju atau tidaknya suatu wilayah atau daerah, bagaimana upaya membangun kualitas hidup manusia sehingga dapat menentukan tingkat kualitas sumber daya manusia suatu wilayah atau daerah.

IPM diperkenalkan oleh United Nations Development Programme(UNDP) pada tahun 1990 dan dipublikasikan secara berkala dalam laporan tahunan Human Development Report (HDR). Dalam publikasi tersebut pembangunan manusia didefinisikan sebagai “a process of enlarging people’s choices” atau proses yang meningkatkan aspek kehidupan masyarakat. IPM yang disusun UNDP memuat tiga indikator, yaitu kesehatan, pendidikan, dan pendapatan, yang menurut UNDP dapat menunjukkan tingkat pembangunan manusia yang sehat dan berumur panjang, berpendidikan dan berketerampilan, serta mempunyai pendapatan untuk dapat hidup layak (UNDP, 1997: 5)

Bagi Indonesia, Indeks Pembangunan Manusia berperan dalam mengidentifikasi tingkatan dalam bidang sosial dan ekonomi, sebagai tinjauan bagi pemerintah dalam mengambil keputusan, dan menjadi tingkat penilaian kinerja pemerintah dalam pembangunan sosial dan ekonomi negara. Selain itu, IPM juga digunakan sebagai salah satu alokator penentuan dana alokasi umum (DAU).Pembangunan manusia di Indonesia dari tahun ke tahun terus mengalami kemajuan, tak terkecuali wilayah provinsi yang berada di Pulau Jawa secara umum selalu memperlihatkan tren meningkat secara linier. Salah satunya ialah Provinsi Jawa Barat yang secara administratif, terbagi menjadi 18 Kabupaten dan 9 Kota. Skor IPM Jawa Barat tidak terlalu berbeda jika dibandingkan dengan IPM nasional. Untuk kedudukan skornya sendiri, Provinsi Jawa Barat belum menunjukkan hasil yang menggembirakan, karena skor IPM-nya tidak lebih baik jika dibandingkan dengan rata -rata IPM nasional. Namun, kenaikan skor IPM Provinsi Jawa Barat mulai terlihat saat periode 2010 hingga 2020, selama periode waktu tersebut indeks pembangunan manusia Provinsi Jawa Barat mengalami peningkatan. Peningkatan tersebut bisa dilihat dari IPM pada tahun 2010 sebesar 66,15 dan pada tahun 2020 menjadi 72,09. Selama periode tersebut, Provinsi Jawa Barat mengalami pertumbuhan IPM rata-rata sebesar 0.86 persen per tahun. Sebelumnya, Kepala Bappeda Jawa Barat, menilai bahwa IPM Jawa Barat di tahun 2020 diperkirakan tidak akan bergerak terlalu jauh karena kondisi pandemi covid-19 yang terjadi saat ini. Namun nyatanya, Indeks Pembangunan Manusia (IPM) pada tahun 2020 menjadi 72,09 atau mengalami kenaikan sebesar 0,06 poin dibandingkan dengan tahun 2019 yang mencapai sebesar 72,03. Besarnya tingkat IPM tersebut menempatkan Provinsi Jawa Barat berada pada status tinggi dengan peringkat ke-10 pemilik tingkat IPM tertinggi secara global. Hal ini menunjukkan kualitas sumber daya manusia di Provinsi Jawa Barat tetap konsisten mengalami kemajuan setiap tahunnya. Dari 27 kabupaten/kota yang berada di Provinsi Jawa Barat, 14 kabupaten/kota mengalami penurunan tingkat IPM pada tahun 2020, salah satunya yaitu Kabupaten Garut dan Kabupaten Karawang. Berkebalikan dengan Kabupaten Garut dan Kabupaten Karawang, Kabupaten Tasikmalaya beserta 12 kabupaten/kota lainnya mengalami peningkatan tingkat IPM di tahun 2020, dimana Kota Bandung, Kota Bekasi, dan Kota Depok memiliki status IPM “sangat tinggi”, sedangkan 24 kabupaten/kota lainnya berstatus “sedang” dan “tinggi”.

Beragamnya status pembangunan manusia tiap kota/kabupaten menuntut pemerintah untuk menyusun strategi yang tepat dalam meningkatkan kualitas sumber daya manusia serta pembangunan nasional. Maka dari itu, pemerintah setempat membutuhkan gambaran kondisi sosial ekonomi kabupaten/kota di wilayah Jawa Barat. Hal tersebut bisa dilakukan dengan mengelompokkan kabupaten/kota di wilayah Jawa Barat untuk mengetahui karakteristik setiap kabupaten/kota berdasarkan Indeks Pembangunan Manusia. Analisis statistika yang dapat digunakan untuk pengelompokkan variabel atau objek adalah Analisis Klaster Hierarki.

Seiring berkembangnya teknologi informasi, hal-hal tersebut dapat diatasi dengan menggunakan Layanan-layanan machine learning. Seperti Microsoft Azure Machine Learning yang merupakan salah satu portal penyedia layanan-layanan komputasi milik Microsoft. Yang nantinya bisa mengelompokan klaster-klaster tersebut dengan menggunakan Machine Learning beserta fitur-fitur pendukungnya seperti K-Means Clustering.

Berdasarkan uraian latar belakang tersebut, ide capstone project ini bertujuan untuk memperoleh pengelompokkan yang objektif dari Kabupaten/Kota di Provinsi Jawa Barat menurut IPM serta mengetahui karakteristik setiap klaster IPM pada tahun 2020. Diharapkan, melalui capstone project ini, pemerintah setempat dapat mengetahui Kabupaten/Kota yang perlu diprioritaskan sehingga program pemerintah dapat tepat sasaran.

# PERUMUSAN MASALAH

﻿ Berdasarkan uraian latar belakang yang ada, yang menjadi inti masalah dalam Capstone Project ini adalah Pemerintah membutuhkan gambaran kondisi social ekonomi kabupaten/kota di wilayah Jawa Barat. Dikarenakan sangat beragamnya kualitas sumber daya manusianya serta pembangunan manusia di tiap kota/kabupaten yang kadang meningkat dan kadang pula menurun, yang secara tidak langsung mengharuskan para pemerintah untuk menyusun strategi yang tepat dalam meningkatkan kualitas sumber daya manusia serta pemerataan pembangunan sumber dayanya.

# SPESIFIKASI KEBUTUHAN

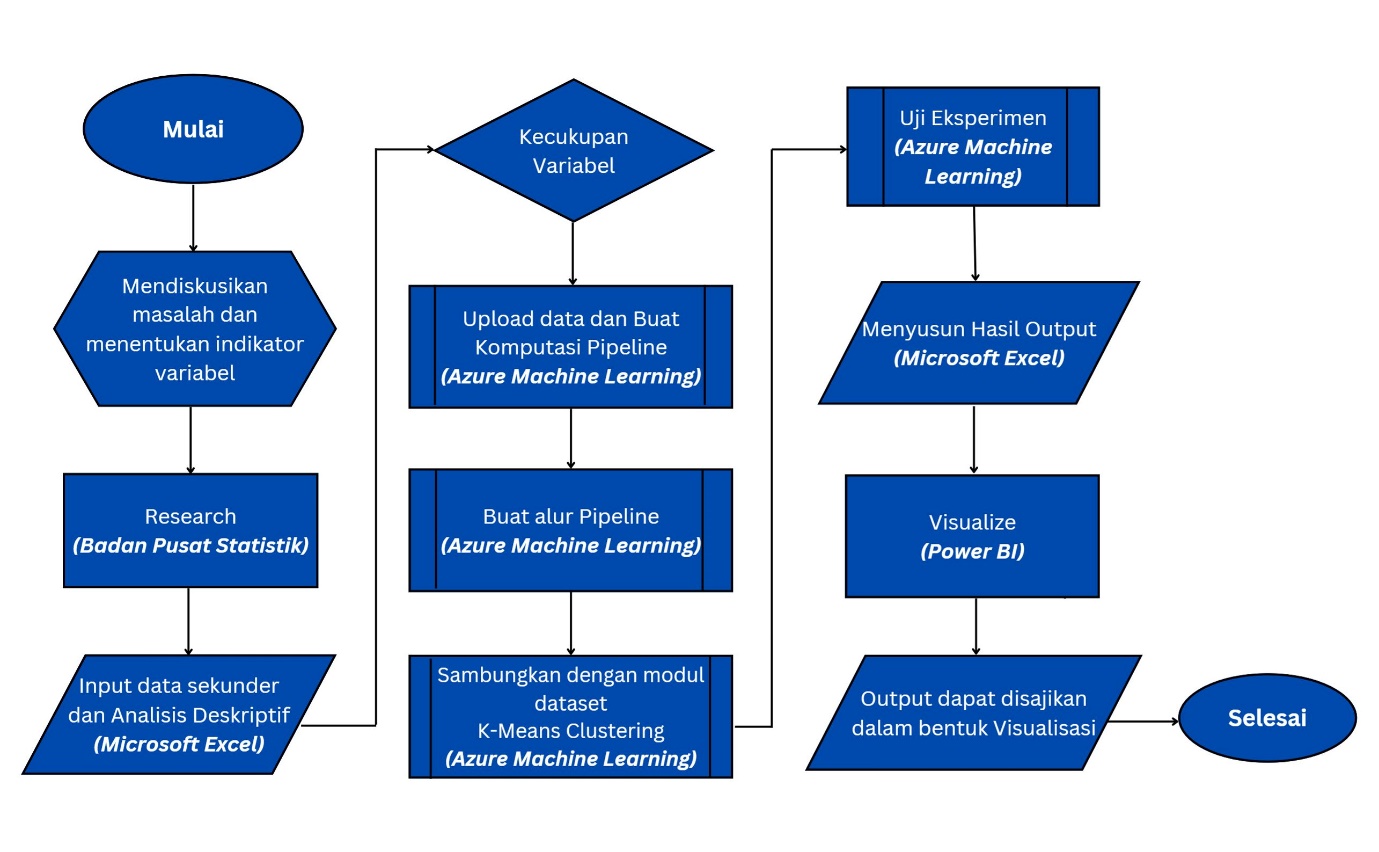
Jika dilihat dari Latar Belakang dan yang menjadi inti masalah dalam Capstone project ini, maka diperlukan untuk melihat Indikator-indikator pembangunan manusia (IPM) yaitu Angka Harapan Hidup dari daerah tersebut, Harapan lama sekolah dari SDM tersebut, Rata-ratalama sekolah di daerah tersebut, dan seberapa banyak pengeluaran perkapita di daerah tersebut ditahun 2020 lalu. Yang dimana data-data tersebut dapat diambil dari Badan Pusat Statistik (BPS).

Disamping itu, untuk mendukung peng-klusterisasian nya, maka diperlukan layanan-layanan pendukung peng-klusterisasian. Seperti Layanan Azure Machine Learning yang akan digunakan, dengan menggunakan fitur pendukungnya yaitu K-Means Clustering yang dapat mengklusterikan menjadi beberapa cluster center.

Dan berdasarkan latar belakang yang ada, pemerintah ini membutuhkan gambaran social yang mudah dipahami, maka diperlukan juga sebuah aplikasi peng-visualisasian seperti Power BI Dekstop untuk memviualisasikan hasil akhir dari sebuah pelaporan peng-klusterisasian yang telah dilakukan. Sehingga, dengan adanya visualisasi tersebut dapat memudahkan pemerintah untuk melihat kondisi social di daerah Kabupaten/Kota di wilayah Jawa Barat.

# RANCANGAN SOLUSI

Proses dan perancangan solusi ini telah melewati beberapa tahapan yang diantara lain sebagai berikut :



# HASIL DAN PEMBAHASAN

Berdasarkan dari permasalahan yang ada, kami menawarkan beberapa solusi untuk mengatasI permasalahan tersebut. Bahwasannya dalam permasalahan tersebut bisa diatasi dengan menggunakan beberapa machine pendukung seperti dibawah ini :

1. ***Microsoft Azure/ Azure Machine Learning Studio***

Microsoft Azure ini adalah platform cloud public milik Microsoft. Yang menyediakan beberapa rangkaian layanan computing, analytic, storage dan jaringan. Sudah dapat diyakini bahwa Azure ini memiliki banyak layanan, seperti Machine Learning, Computer Vision, Cognitive Service, dsb. Maka, dengan adanya layanan tersebut, kita bisa memanfaatkan layanan tersebut untuk menjadikan itu sebuah solusi yang kita gunakan dalam permasalahan ini.

Salah satunya dengan menggunakan Azure Machine Learning, yang dimana layanan tersebut dapat membantu menyederhanakan beberapa tugas yang diperlukan untuk menyiapkan data, melatih model, dan menerapkan layanan prediktif. Didalam Azure machine learning ini masih terdapat beberapa lagi jenis machine learning seperti Klasifikasi dan Clustering. Yang dimana, salah satu jenis tersebut bisa kita gunakan untuk meng-klusterisasikan yaitu fitur Clustering.

Clustering ini adalah Teknik unsupervised machine learning tanpa pengawasan yang digunakan untuk mengelompokan entitas serupa berdasarkan fiturnya. Dengan menggunakan jenis Clustering Machince Learning, didalamnya pun masih terdapat beberapa fitur-fitur pendukung unntuk peng-klusterisasian, salah satunya adalah K-Means Clustering. Yang dimana K-Means Clustering ini digunakan untuk clustering dokumen. Prinsip utama dari K-Means Clustering inin sendiri adalah menyusun pusat massa (centeroid) dari sekumpulan data. Sehingga sebelum diterapkan proses algoritma K-Means Clustering, dokumen akan diproses terlebih dahulu. Sedangkan tujuan dari K-Means Clustering ini adalah untuk meminimalisasikan fungsi objective yang telah diset dalam proses clustering tadi. Sehingga, variasi data yang ada di cluster lainnya dapat dimaksimalkan.

1. ***Microsoft Excel***

Microsoft Excel adalah salah satu program aplikasi lembar kerja spreadsheet yang dibuat dan didistribusikan oleh Microsoft Corporation untuk mengolah data dalam berbagai bentuk, seperti rumus, perhitungan dasar, pembuatan table. Grafik hingga pengelolaan data. Alasan kami menyantumkan aplikasi ini sebagai sebuah solusi adalah untuk mengolah hasil dan memilah hasil yang didapati dari pengklusterisasian sebelumnya. Sehingga dengan menggunakan Microsoft Excel ini dapat memudahkan kita untuk memvisualisasikan hasil pelaporannya.

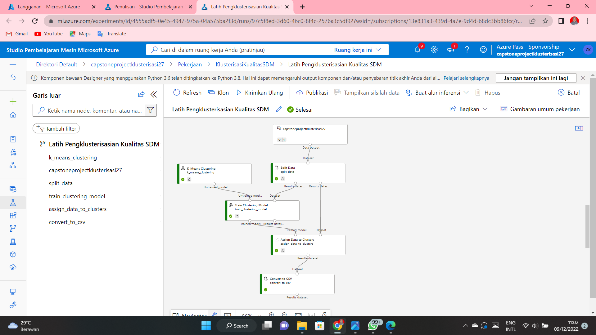
1. ***Microsoft Power BI***

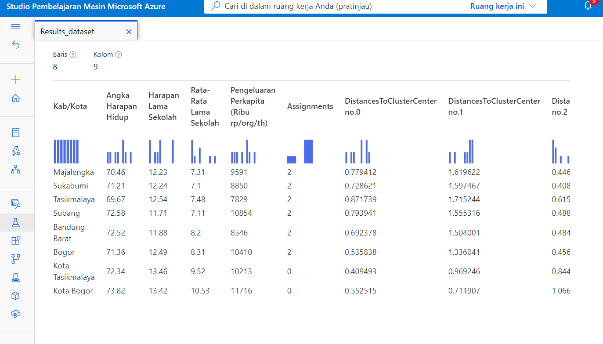
Power BI adalah kumpulan lengkap layanan dan alat bantu yang dapat memungkinkan kita untuk tersambung, mengubah dan memvisualisasikan data. Dengan Power BI ini, kita dapat menyambungkan ke beberapa sumber data yang berbeda dan menggabungkannya adalah ke model data.

Didalam Power BI ini terdapat banyak sekali tools-tools untuk memvisualisasikan data yang kita miliki. Sehingga dengan menggunakan tools-tools yang ada di Power BI, kita nantinya akan dengan mudah melihat hasil akhir dari pengklusterisasian yang kita gunakan. Seperti daerah-daerah apa saja yang indicator kualitas SDM nya lebih unggul, dan daerah-daerah apa saja yang perlu pemerataan.

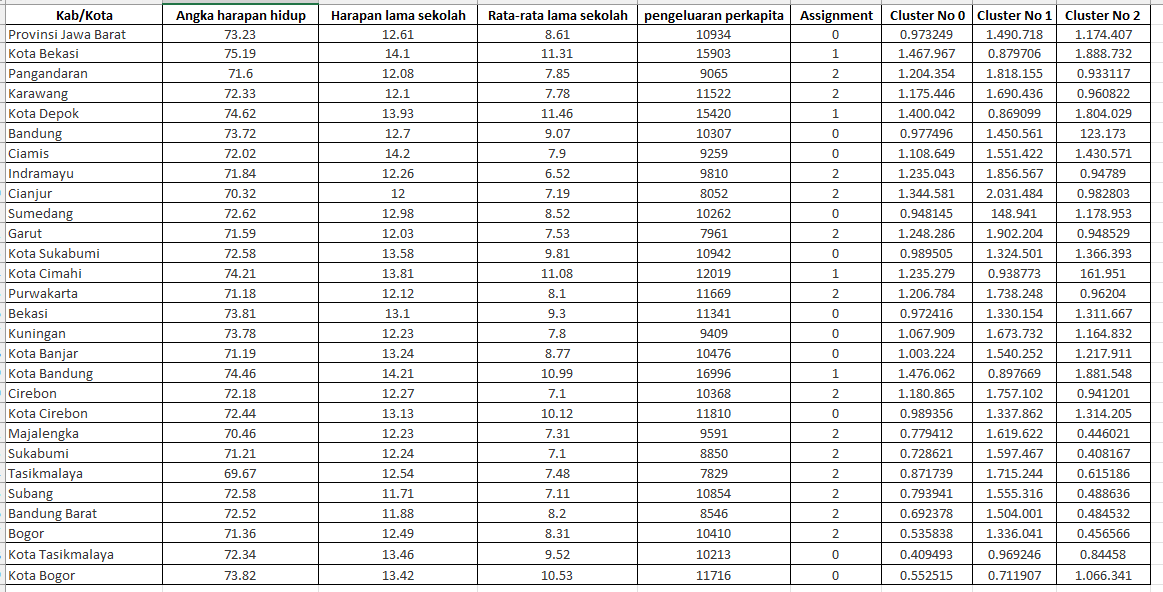
Akan tetapi solusi yang kami usulkan ini memiliki keterbatasan, yang dimana car aini tidak bisa menghubungkan data yang diambil secara real-time. Maksudnya, cara ini tidak bisa secara otomatis menghubungkan data dari BPS ke machine learning ini. Dikarnakan disini kami tidak menghubungkan data set tersebut kedalam uji eksperimen pada saat pengujian. Disisi lain jika dibandingkan dengan cara-cara yang sudah ada seperti menggunakan metode perhitungan statistic dengan alat pendukungnya yaitu aplikasi SPSS, cara ini lebih mudah untuk dilakukan. Terlebih lagi jika solusi yang kami usulkan ini menerapkan atau mengaktifkan alur inferensi secara realtime, yang memudahkan user mengambil data klusterisasinya. Jika menggunakan perhitungan SPSS, kita perlu menginput angka-angkanya terlebih dahulu ke dalam rumus-rumus yang bisa digunakan. Tetapi, jika menggunakan layanan Azure Machine Learning ini kita hanya perlu meng-upload datanya dan membuat komputasi, kemudian dikarenakan solusi ini mengadopsi system Drag and Drop, maka akan memudahkan kita dalam penghubungan klusterisasi.

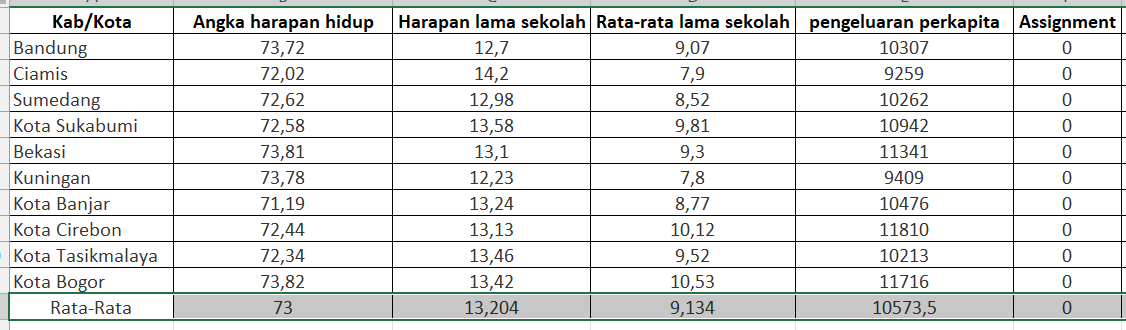
Berikut adalah pipeline pengukuran dan hasil kluster dengan menggunakan Azure Machine Learning yang dihubungkan dengan fitur pendukungnya, yaitu K-Means Clustering.

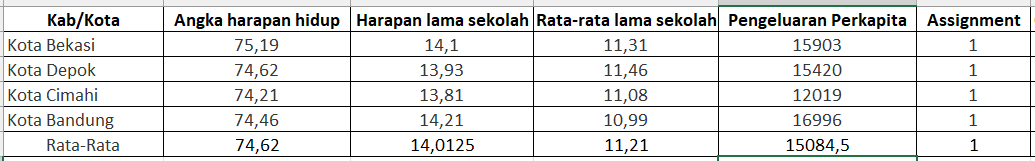


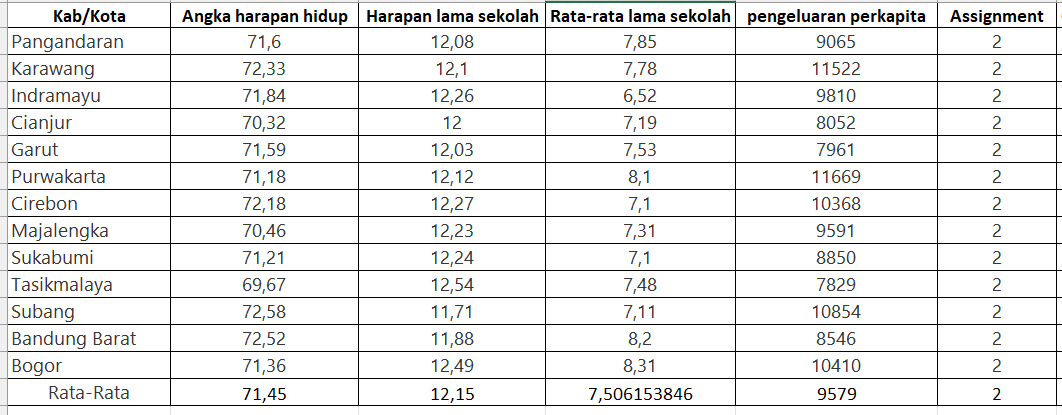


Dan berikut adalah *Review* Output yang dikeluarkan oleh Machine Learning

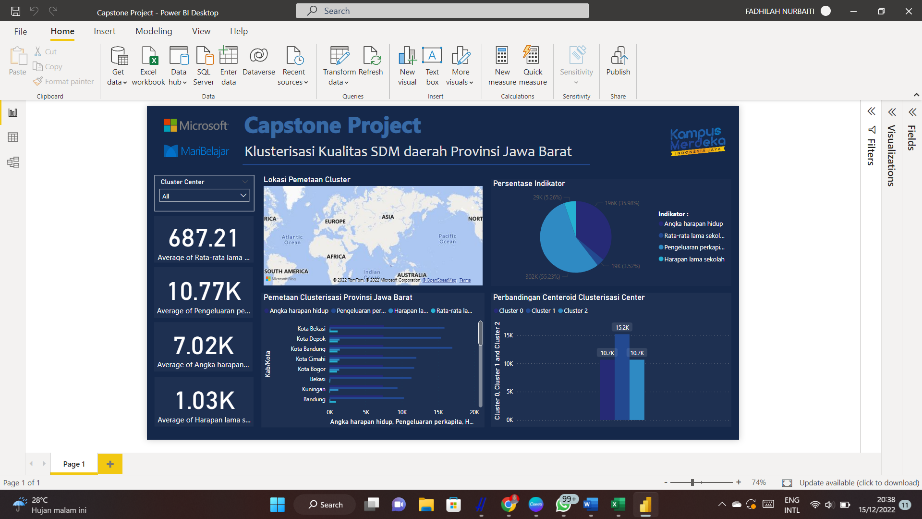


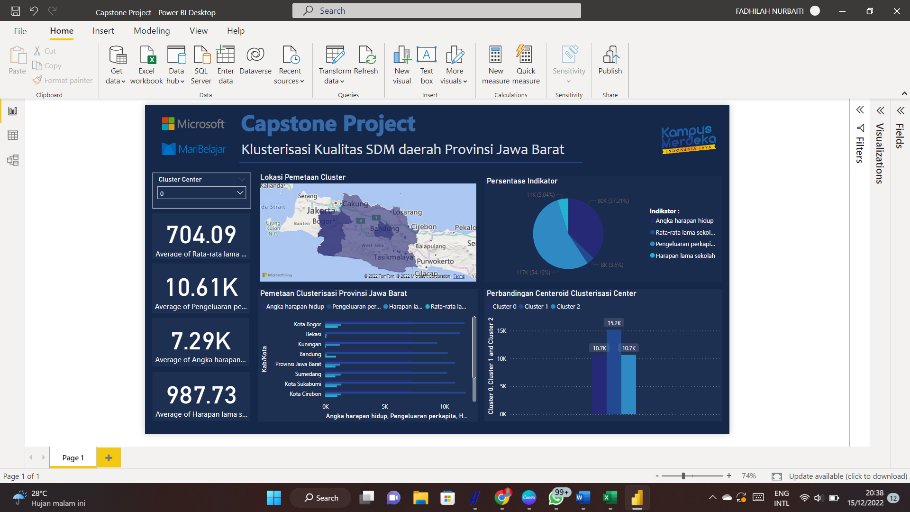


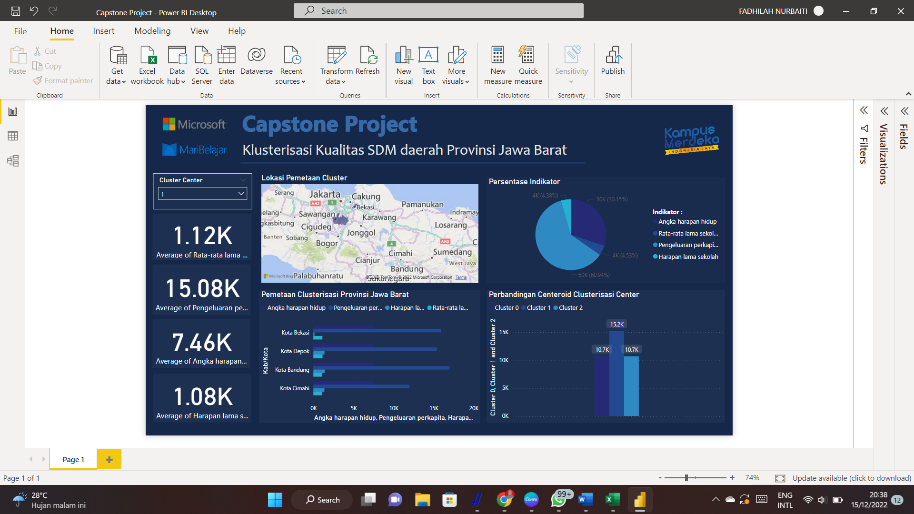


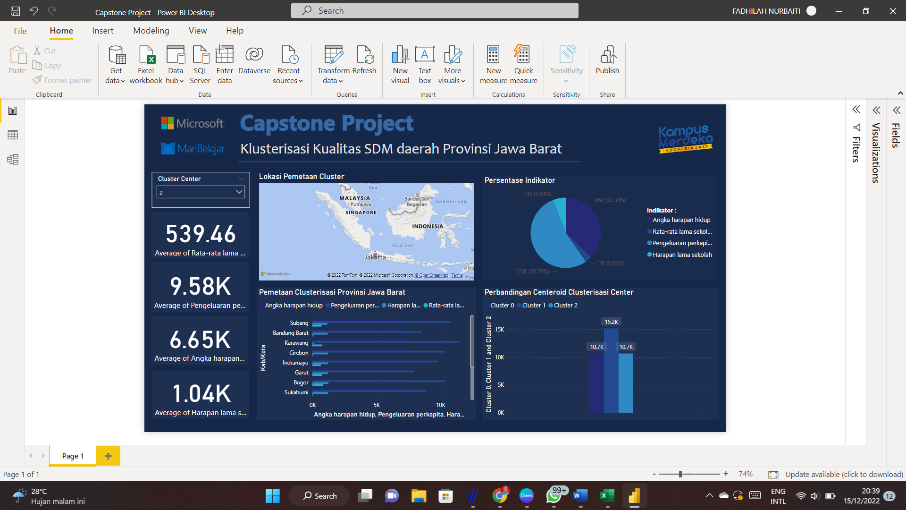


Kemudian berikut adalah visualisasi hasil akhirnya









Dapat dilihat dari visualisasi akhir diatas dan berdasarkan hasil uji eksperimen yang dilakukan di Azure Machine Learning Studio, bahwasannya kategori Kualitas Sumber Daya Manusia ini akhirnya terpecah menajdi 3 Cluster Center, yaitu sebagai berikut :

1. Cluster Center No. 0 cenderung memiliki kualitas sumber daya manusia yang “sedang” jika dibandingkan dengan cluster lainnya. Pada cluster ini juga mempunyai tingkat Angka Harapan Hidup, Rata-Rata Lama Sekolah, Harapan Lama Sekolah yang “sedang” dengan Pengeluaran Perkapita pada tingkat “rendah”.
2. Cluster Center No. 1 cenderung memiliki kualitas sumber daya manusia yang “tinggi” jika dibandingkan dengan kedua cluster lainnya. Pada cluster ini juga mempunya tingkat Angka Harapan Hidup, Rata-Rata Lama Sekolah, Harapan Lama Sekolah yang “tinggu” dengan Pengeluaran Perkapita pada tingkat sama-sama “tinggi”.
3. Cluster Center No. 2 cenderung memiliki kualiras sumber daya manusia yang “rendah” jika dibandingkan dengan Cluster lainnya. Pada Cluster ini juga mempunya tingkat Angka Harapan Hidup, Rata-Rata Lama Sekolah, Harapan Lama Sekolah yang “rendah” dengan Pengeluaran Perkapita pada tingkat sama-sama “rendah”

# KESIMPULAN

Sehingga dari output tersebut dapat disimpulkan bahwa kota/kabupaten di Jawa Barat ini terbagi mejadi 3 Cluster Center dimana Kluster yang memiliki status IPM “sangat tinggi”adalah Kluster 1 yaitu Kota Bekasi, Kota Depok, Kota Cimahi dan Kota Bandung. Sedangkan kluster dengan status IPM “terendah” adalah kluster 2 yang berjumlah 12 Kabupaten/kota.

Dari peng-klusterisasian yang telah dilakukan, dapat diketahui bahwa gambaran IPM Kota/Kabupaten di Jawa Barat Tahun 2020 masih belum merata . Sehingga, diharapkan pemerintah Provinsi Jawa Barat dapat memprioritaskan 12 Kabupaten dengan IPM yang tergolong rendah yaitu Pangandaran, Kabupaten Karawang, Indramayu, Cianjur, Garut, Purwakarta, Cirebon, Majalengka, Sukabumi, Tasikmalaya, Subang, Bandung Barat, Bogor, supaya dapat diberikan penanganan yang tepat sehingga masalah rendahnya tingkat IPM dapat teratasi, karna dapat mempengaruhi tingkat kualitas sumber daya manusia didalam daerah tersebut.

# LAMPIRAN

Berikut penulis lampirkan terkait hasil dari Klusterisasi Kualitas SDM (Menggunakan K-Means Clustering):

1. Link Github

<https://github.com/ninditaayu/Capstone-Project.git>

1. Link Pitching Video

<https://youtu.be/4cJ5ftackxM>

1. Link Demonstrasi Video Solusi

<https://youtu.be/ZLLLSevaWf4>