

Analisis Clustering Provinsi di Indonesia berdasarkan Karakteristik Kesehatan

Menggunakan Metode *Principal Component Analysis*
dan *K Means Clustering*



Anggota Kelompok



Hafez Afghan

06211940000018
Statistika FSAD ITS 2019



Sayyid Nur C

06211940000089
Statistika FSAD ITS 2019

Table of contents

01.

Pendahuluan

Latar belakang, masalah dan tujuan, manfaat.

03.

Analisis

Analisis dan Pembahasan.

02.

Metodologi

Sumber data, variabel penelitian, dan langkah penelitian.

04.

Kesimpulan & Saran

Kesimpulan dari hasil analisis. Saran kepada stakeholder terkait.





01.

Pendahuluan

Latar Belakang



Bidang kesehatan di Indonesia perlu diperhatikan lebih, dimana masih banyaknya kasus - kasus kematian, dan kurangnya fasilitas kesehatan



Banyaknya provinsi di Indonesia menyebabkan kebijakan yang diberikan harus berbeda - beda tergantung karakteristiknya



Diperlukan *clustering* Provinsi di Indonesia berdasarkan karakteristiknya untuk memberikan kebijakan bidang kesehatan dengan baik dan tepat sasaran

Rumusan Permasalahan

Bagaimana mengimplementasikan PCA dan *k means clustering* dalam mengelompokkan Provinsi di Indonesia berdasarkan karakteristik kesehatan setiap Provinsi.

Tujuan Penelitian

1

Membuat sebuah model *clustering* Provinsi di Indonesia untuk diberikan kepada pemangku kebijakan di bidang kesehatan.

2

Menerapkan PCA dan *K Means Clustering* dalam pengelompokkan Provinsi di Indonesia berdasarkan karakteristik kesehatan setiap Provinsi.

Manfaat

Manfaat dari penelitian ini adalah diperoleh sebuah model *clustering* Provinsi di Indonesia berdasarkan karakteristik kesehatan setiap Provinsi sehingga dapat menjadi acuan bagi pemerintah dalam memberikan kebijakan mengenai kesehatan pada sebuah Provinsi.

02. Metodologi Penelitian



Sumber Data

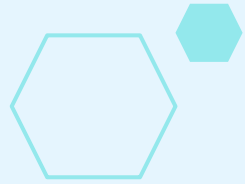
Data yang diteliti merupakan data sekunder yang diambil dari situs resmi Profil Kesehatan Indonesia. Data yang diambil berdasarkan Provinsi di Indonesia pada tahun 2019.



Variabel Penelitian

Nama Variabel	Keterangan	Skala Data
X1	Kepadatan Penduduk	Rasio
X2	Jumlah Tenaga Medis	Rasio
X3	Jumlah Puskesmas	Rasio
X4	Persentase Rumah Tangga yang Memiliki Air Minum Layak	Rasio
X5	Jumlah Kasus DBD	Rasio
X6	Jumlah Kematian Ibu	Rasio
X7	Persentase Desa yang Memanfaatkan Dana Desa untuk Kesehatan	Rasio

Langkah Penelitian



Mengumpulkan
Data Sekunder

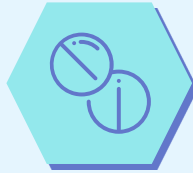
2

Analisis *Clustering*
Pertama (*K-Means*)

4



1



3



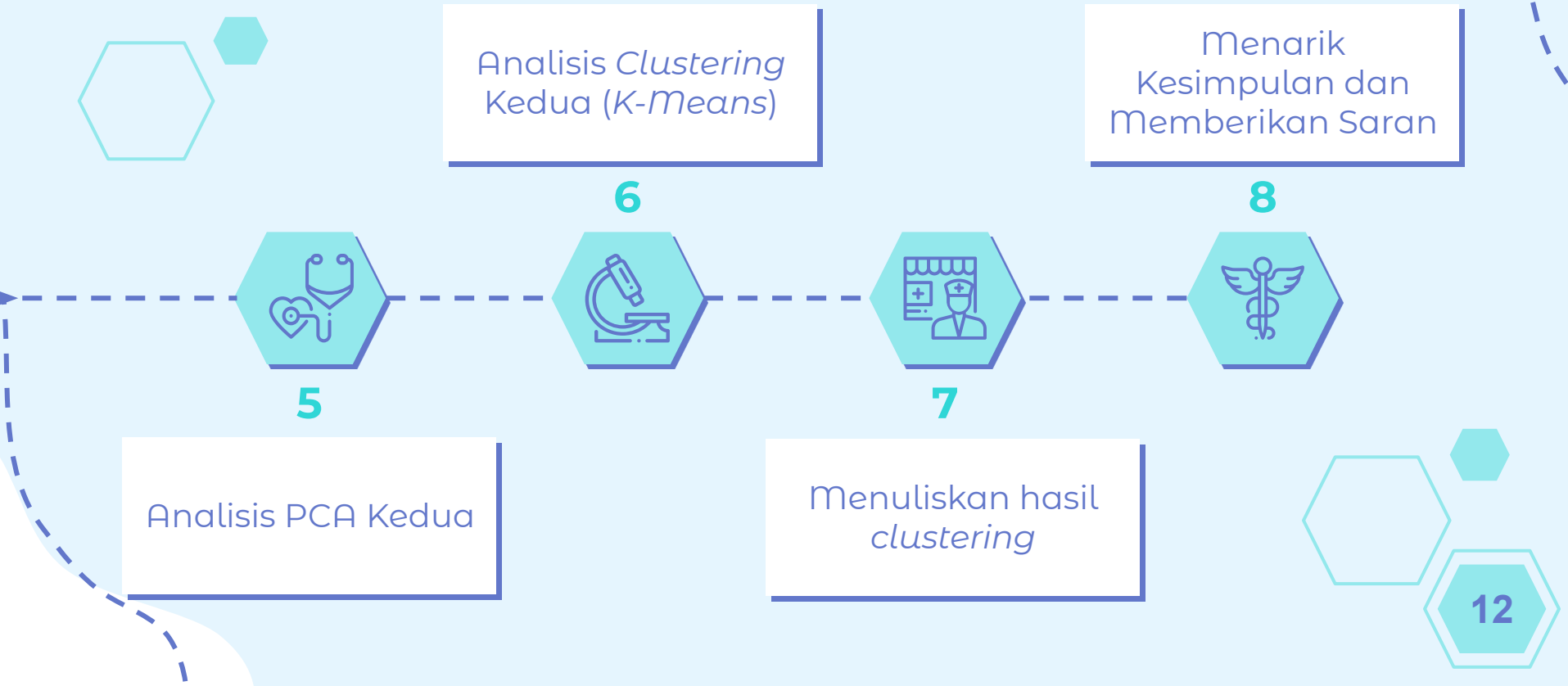
Merumuskan
Permasalahan

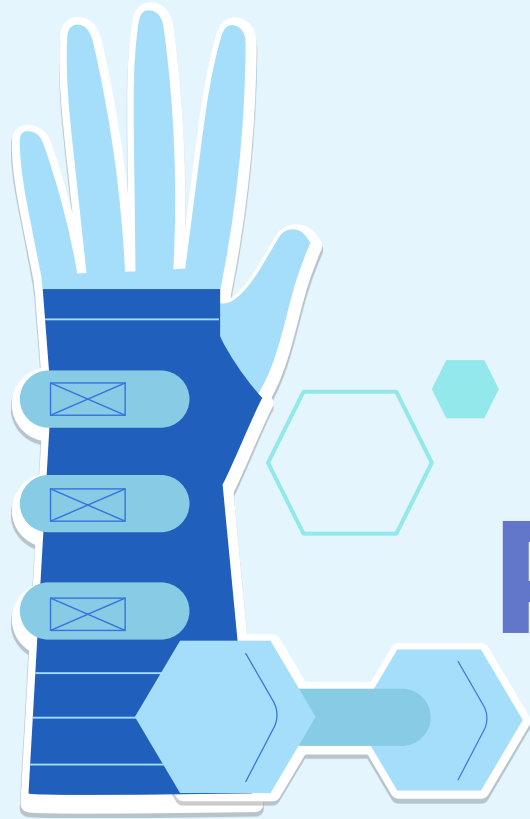
Analisis PCA
Pertama



11

Langkah Penelitian



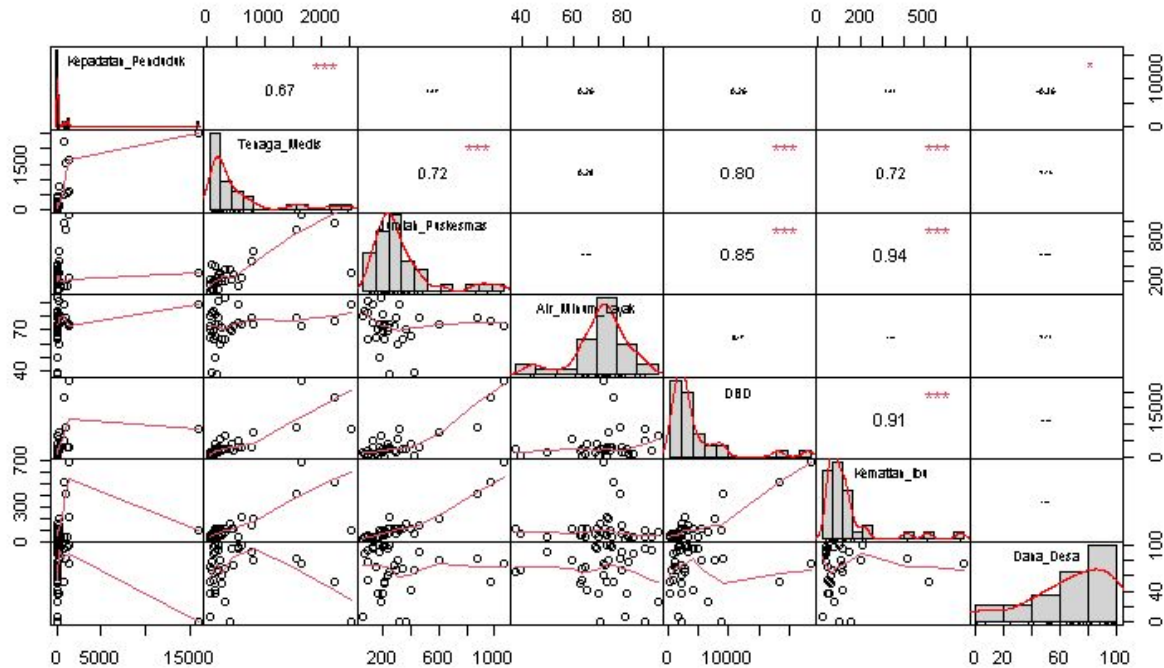


03. Analisis & Pembahasan

Statistika Deskriptif

Kriteria	X1	X2	X3	X4	X5	X6	X7
Minimum	9,32	55	40	37,37	236	21	0
Kuartil 1	54,8	116,5	161,5	66,41	1513	47,5	51,82
Median	103,42	194,5	218,5	73,42	2348	76,5	73,43
Mean	742,01	461,2	298,1	72,14	4063	124,1	67,2
Kuartil 3	261,23	515,5	334,5	79,32	4109	117,8	91,99
Maksimum	15900,08	2522	1072	93,77	23483	684	100

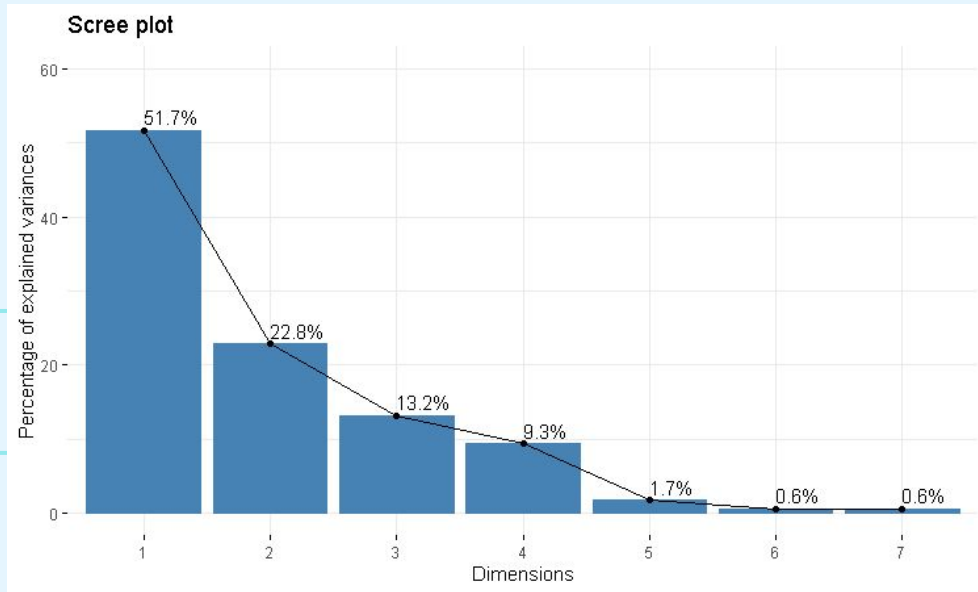
PCA Pertama



Korelasi dan Distribusi Data

Dilihat dari pola distribusi 7 variabel disamping menunjukkan tidak ada yang berdistribusi normal. Berdasarkan korelasi juga menunjukkan korelasi yang tinggi untuk variabel seperti jumlah puskesmas dan jumlah kematian ibu.

PCA Pertama

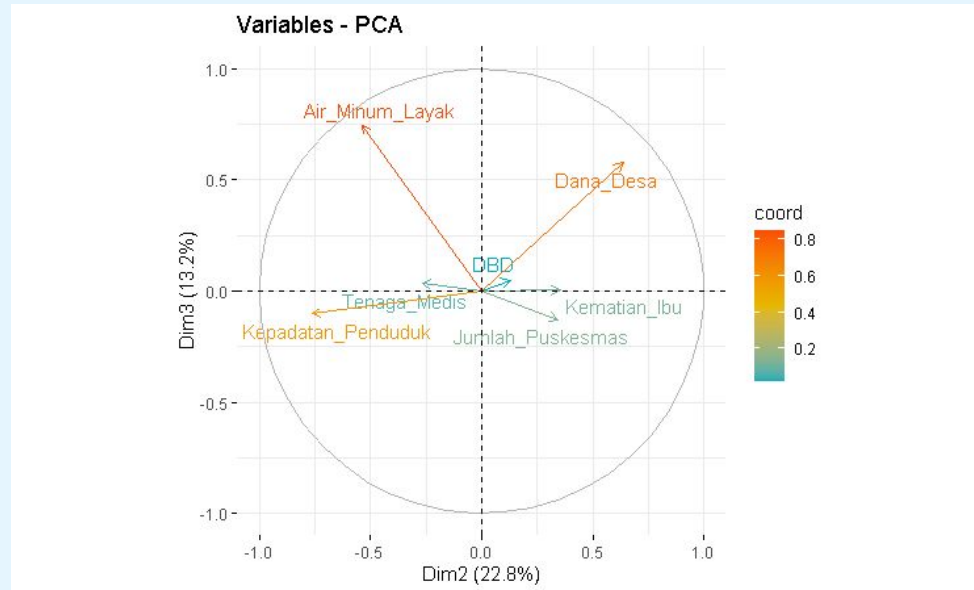


	PC 1	PC 2	PC 3	PC 4	PC 5
Eigen	3,62	1,6	0,92	0,65	0,12

Scree Plot

Grafik disamping menunjukkan bahwa pada PC 2 mampu menjelaskan total 74,5% variansi dari data dengan nilainya yang tidak berbeda jauh dari PC 3. Berdasarkan nilai eigen value > 1 menunjukkan nilai PC 2 dengan nilai 1,6 sehingga disimpulkan jumlah PC optimum adalah 2.

PCA Pertama



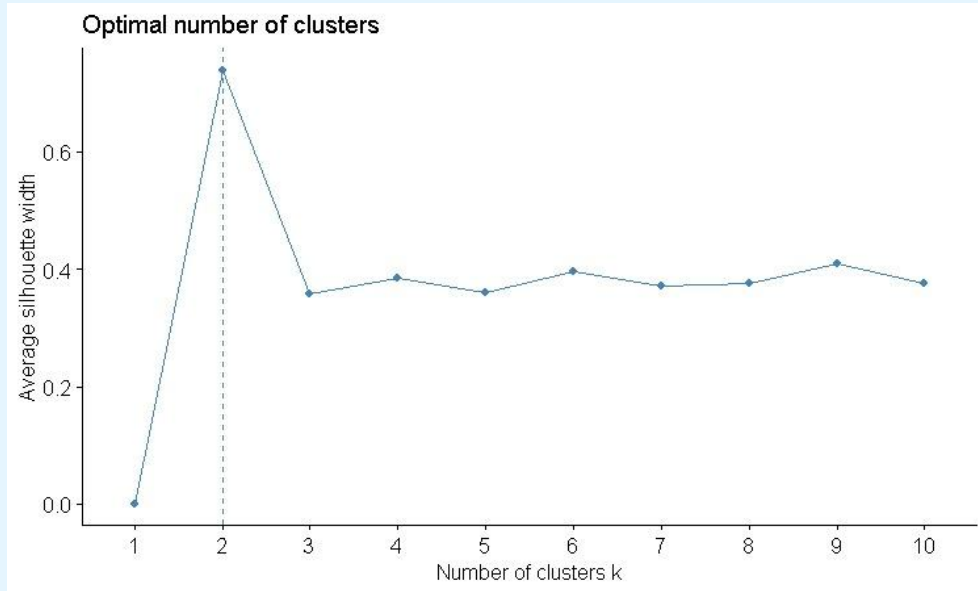
PCA Equation

PC1 = Jumlah Tenaga Medis, Jumlah Puskesmas, Kasus DBD, Kematian Ibu

PC2 = Kepadatan Penduduk, Persentase Rumah Tangga Dengan Akses Air Minum Layak, Persentase Desa Yang Memanfaatkan Dana Desa Untuk Kesehatan

	X1	X2	X3	X4	X5	X6	X7
PC1	0,222	0,485	0,471	0,104	0,498	0,48	-0,065
PC2	-0,601	-0,212	0,272	-0,423	0,101	0,279	0,503

Clustering Pertama

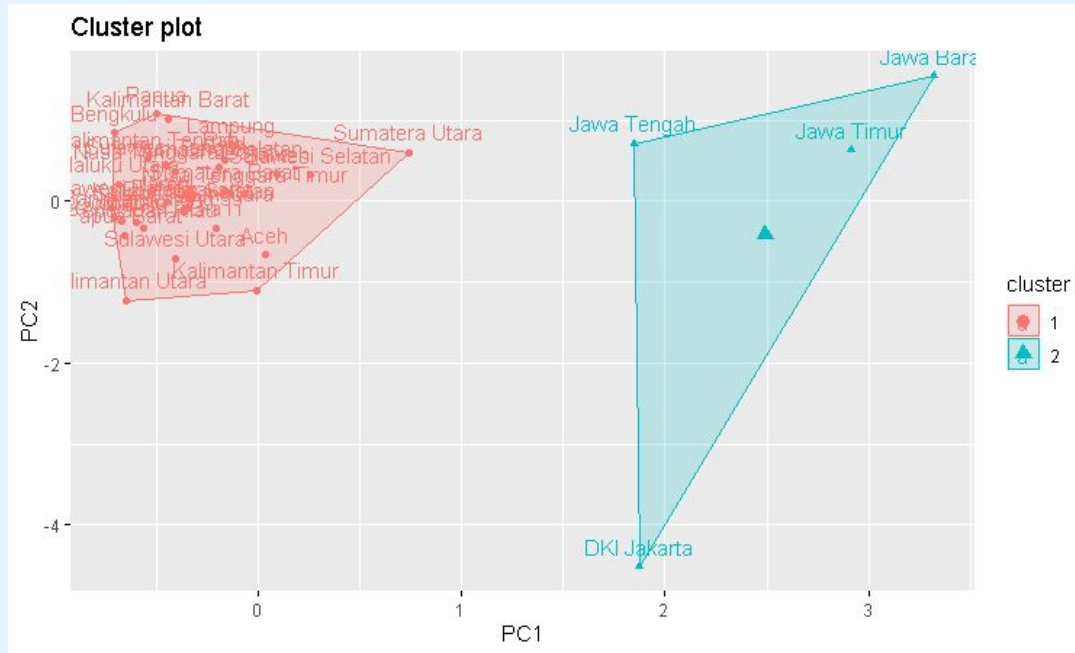


Silhouette Plot

Grafik disamping menunjukkan bahwa ketika jumlah cluster 2 memiliki nilai silhouette paling tinggi. Namun akan dianalisis juga $k = 3, 4$, dan 5. Model Optimum ditentukan melalui nilai F Pseudo yang tertinggi, yaitu $k = 2$, sehingga $k = 2$ merupakan model optimum.


K-means	2 Cluster	3 Cluster	4 Cluster	5 Cluster
F Pseudo	11,487	8,956	3,169	1,772

Clustering Pertama



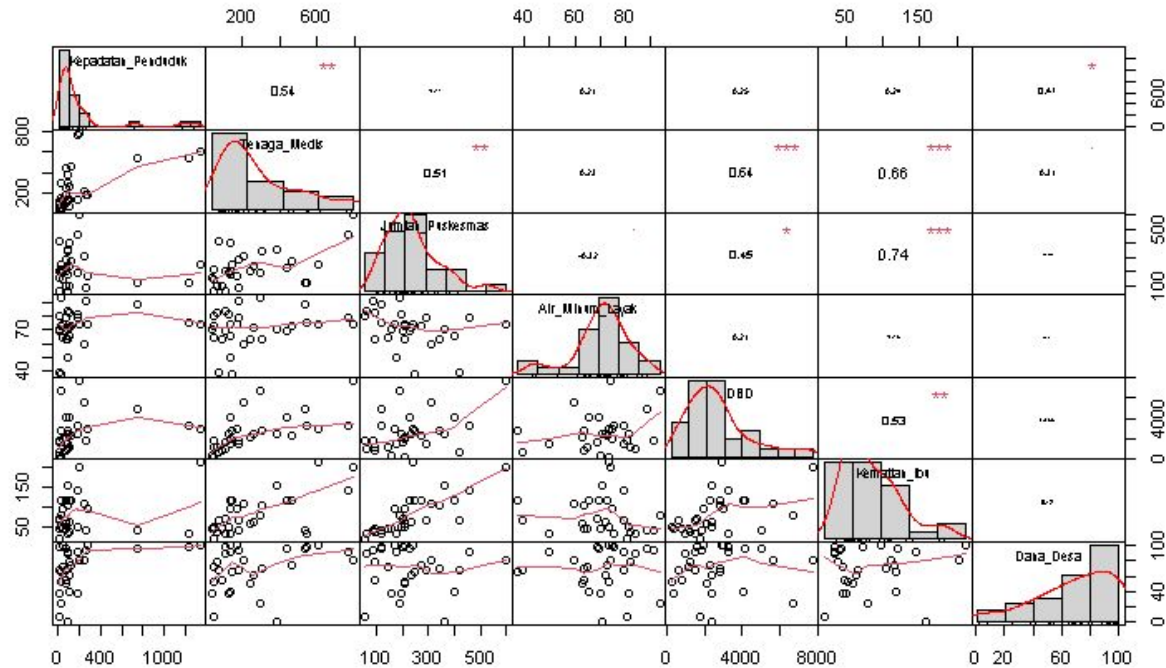
Clustering Plot

Diperoleh bahwa terdapat 2 cluster, dimana cluster 2 terdiri dari Provinsi Jawa Tengah, Jawa Timur, Jawa Barat, dan DKI Jakarta. Sedangkan cluster 1 terdiri dari provinsi lainnya.



Karena sangat banyaknya Provinsi yang dikelompokkan dalam suatu *cluster*, maka akan dianalisis *clustering* tahap dua. Dengan menggunakan PCA tahap dua dalam analisisnya.

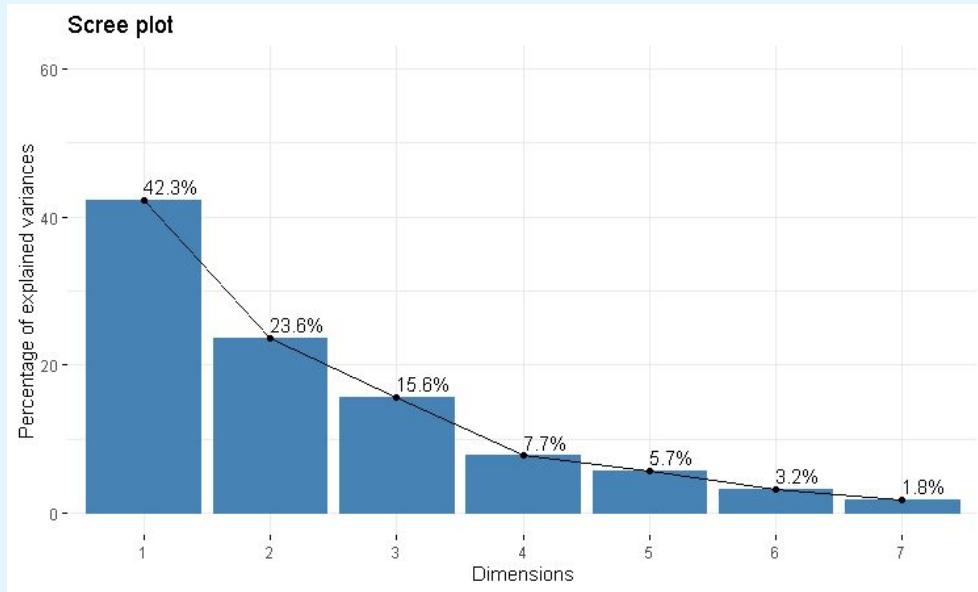
PCA Kedua



Korelasi dan Distribusi Data

Dilihat dari pola distribusi 7 variabel disamping menunjukkan tidak ada yang berdistribusi normal. Berdasarkan korelasi juga menunjukkan tidak adanya korelasi yang tinggi untuk variabel.

PCA Kedua

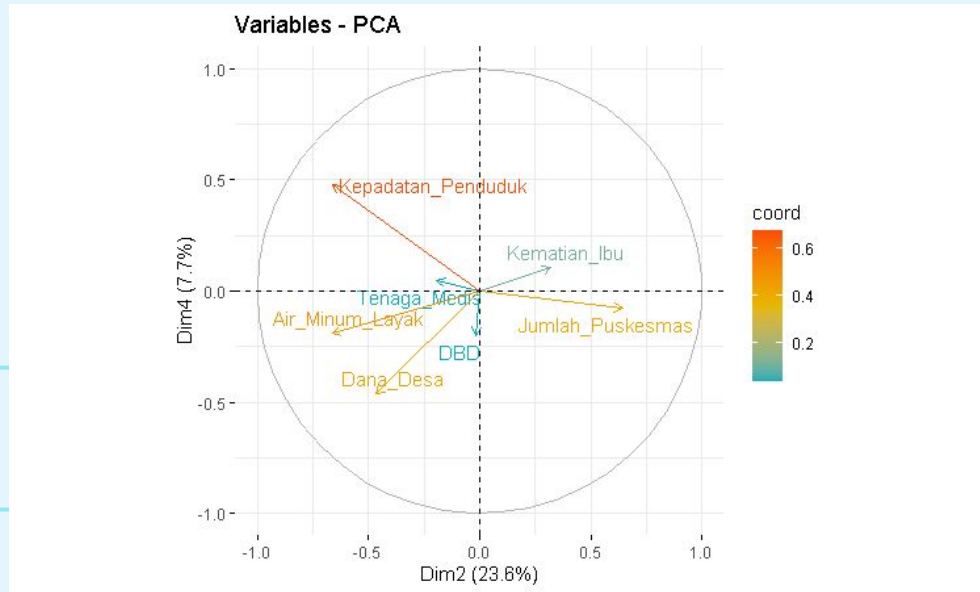


	PC 1	PC 2	PC 3	PC 4	PC 5
Eigen	2,96	1,65	1,09	0,54	0,4

Scree Plot

Grafik disamping menunjukkan bahwa pada PC 3 mampu menjelaskan total 81,5% variansi dari data dengan nilainya yang tidak berbeda jauh dari PC 4. Berdasarkan nilai eigen value > 1 menunjukkan nilai PC 3 dengan nilai 1,09 sehingga disimpulkan jumlah PC optimum adalah 3.

PCA Kedua



PCA Equation

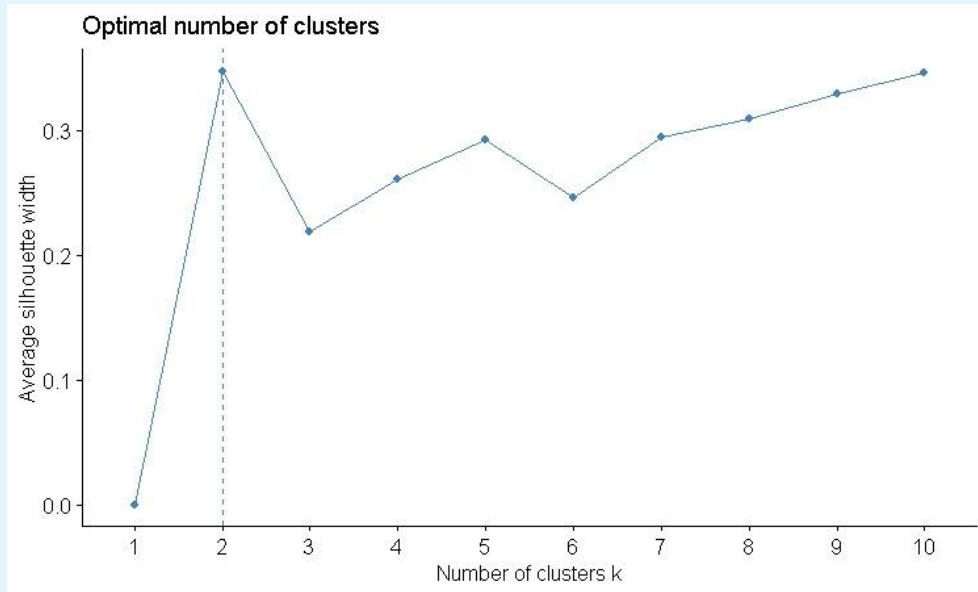
PC1 = Jumlah Tenaga Medis, Kasus DBD, Kematian Ibu

PC2 = Kepadatan Penduduk, Jumlah Puskesmas

PC3 = Presentase Rumah Tangga Dengan Akses Air Minum Layak, Presentase Desa Yang Memanfaatkan Dana Desa Untuk Kesehatan

	X1	X2	X3	X4	X5	X6	X7
PC 1	0,285	0,531	0,4	0,045	0,445	0,492	0,187
PC 2	-0,516	-0,15	0,503	-0,515	-0,013	0,248	-0,361
PC 3	-0,236	0,062	-0,02	0,627	0,35	-0,04	-0,651

Clustering Kedua

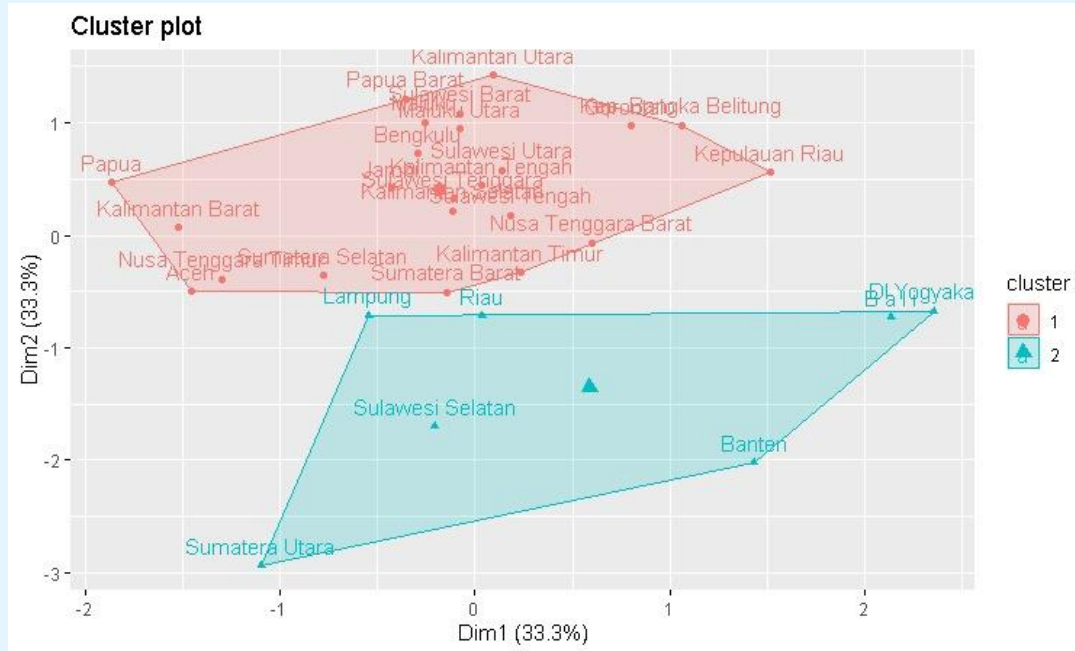


K-means	2 Cluster	3 Cluster	4 Cluster	5 Cluster
F Pseudo	8,991	3,23	1,814	0

Silhouette Plot

Grafik disamping menunjukkan bahwa ketika jumlah cluster 2 memiliki nilai silhouette paling tinggi. Namun akan dianalisis juga $k = 3, 4$, dan 5. Model Optimum ditentukan melalui nilai F Pseudo yang tertinggi, yaitu $k = 2$, sehingga $k = 2$ merupakan model optimum.

Clustering Kedua



Clustering Plot

Diperoleh bahwa terdapat 2 cluster, dimana cluster 2 terdiri dari Provinsi Lampung, Riau, Bali, DI Yogyakarta, Sulawesi Selatan, Banten, Sumatera Utara. Sedangkan cluster 1 terdiri dari provinsi lainnya.

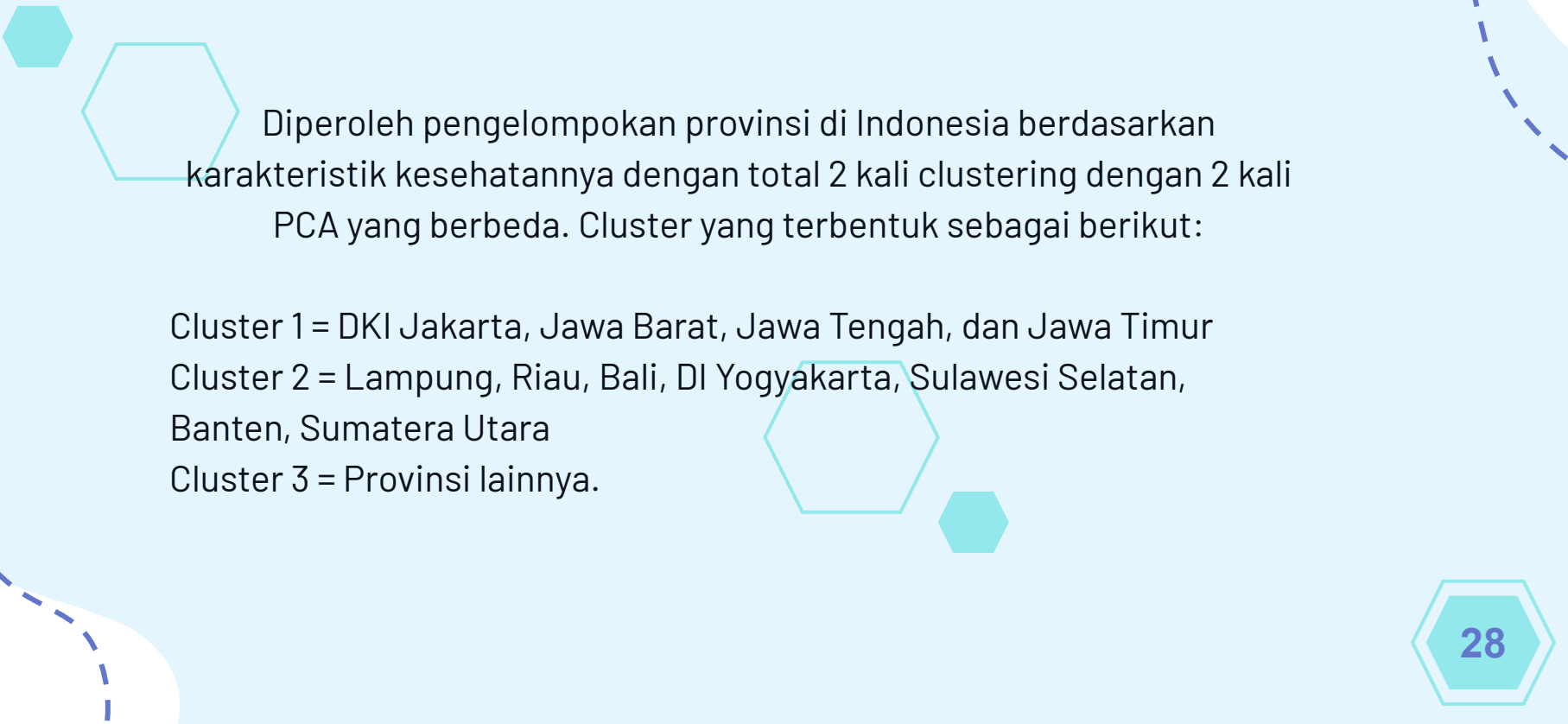
Hasil Clustering

Cluster	Provinsi
Cluster 1	Jawa Tengah, Jawa Timur, Jawa Barat, dan DKI Jakarta
Cluster 2	Lampung, Riau, Bali, DI Yogyakarta, Sulawesi Selatan, Banten, dan Sumatera Utara
Cluster 3	Aceh, Sumatera Barat, Jambi, Sumatera Selatan, Bengkulu, Kep. Bangka Belitung, Kepulauan Riau, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua, dan Papua Barat

Kesimpulan & Saran



Kesimpulan

The slide features a light blue background with decorative elements. On the left, there is a small solid teal hexagon and a larger teal-outlined hexagon. On the right, a dashed blue line curves from the top towards the middle. At the bottom right, there is a teal-outlined hexagon containing the page number. The text is centered and describes the clustering of Indonesian provinces based on health characteristics.

Diperoleh pengelompokan provinsi di Indonesia berdasarkan karakteristik kesehatannya dengan total 2 kali clustering dengan 2 kali PCA yang berbeda. Cluster yang terbentuk sebagai berikut:

Cluster 1 = DKI Jakarta, Jawa Barat, Jawa Tengah, dan Jawa Timur

Cluster 2 = Lampung, Riau, Bali, DI Yogyakarta, Sulawesi Selatan, Banten, Sumatera Utara

Cluster 3 = Provinsi lainnya.

Saran

Hasil clustering menunjukkan perbedaan perbedaan yang mencolok pada karakteristik kesehatan masing-masing provinsi di cluster. Hal ini diharapkan dapat dijadikan acuan untuk pemerataan kesehatan di wilayah-wilayah dengan kesehatan yang rendah, terutama pada cluster 3 yang mencakup 23 provinsi di Indonesia. Kepada peneliti selanjutnya agar dapat menggunakan data yang terkini, dikarenakan per 17 November 2022 provinsi di Indonesia sejumlah 38 provinsi, sehingga perlu analisis *clustering* terhadap provinsi - provinsi baru tersebut.

Thanks

Do you have any questions?

Analisis Data C
Tugas PCA Minggu 12



Introduction

You can give a brief description of the topic you want to talk about here. For example, if you want to talk about Mercury, you can say that it's the smallest planet in the entire Solar System



Our center

Do you know what helps you make your point clear?

Lists like this one:

- They're simple
- You can organize your ideas clearly
- You'll never forget to buy milk!

And the most important thing: the audience won't miss the point of your presentation



Mission and vision



Mission

Mercury is the closest planet to the Sun and the smallest one in the Solar System—it's only a bit larger than the Moon



Vision

Venus has a beautiful name and is the second planet from the Sun. It's terribly hot—even hotter than Mercury



“This is a quote, words full of wisdom
that someone important said and can
make the reader get inspired.”

—Someone Famous

Key numbers

20,000

This is the number of patients we had last month

50,000

This is the number of total visitors we had last month

5,000

This is the number of surgeries performed so far

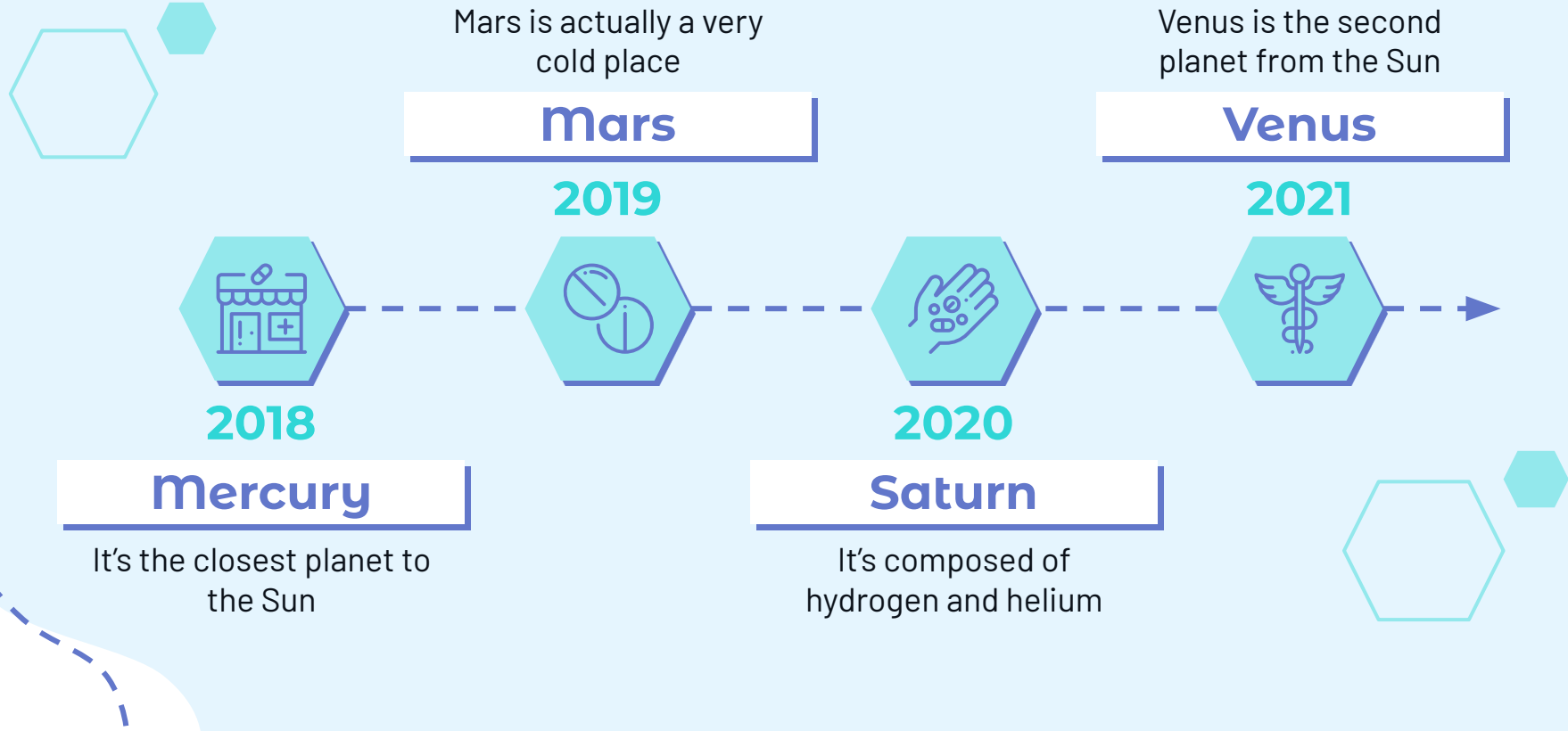




A brief story

Mercury is the closest planet to the Sun and the smallest one in the Solar System—it's only a bit larger than the Moon. The planet's name has nothing to do with the liquid metal, since Mercury was named after the Roman messenger god

milestones reached



Our complete process

Step 1

Mars is actually a very cold place

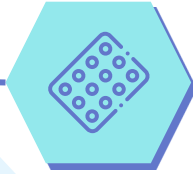


Step 2

Venus is a very hot planet

Step 3

Saturn is the ringed planet



Step 4

Mercury is a very small planet





Patient care

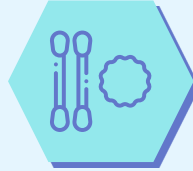
Mercury is the closest planet to the Sun and the smallest one in the Solar System—it's only a bit larger than the Moon

Areas we cover



Venus

Venus is the second planet from the Sun



Jupiter

It's the biggest planet in the Solar System



Mars

Despite being red, Mars is a cold place



Saturn

Saturn is a gas giant and has several rings

These are our main services



Venus

Venus is the second planet from the Sun



Jupiter

It's the biggest planet in the Solar System



Mars

Despite being red, Mars is a cold place



Saturn

Saturn is a gas giant and has several rings

Key accomplishments



Mercury

It's the closest planet to the Sun and the smallest one in the Solar System



Venus

Venus has a beautiful name and is the second planet from the Sun



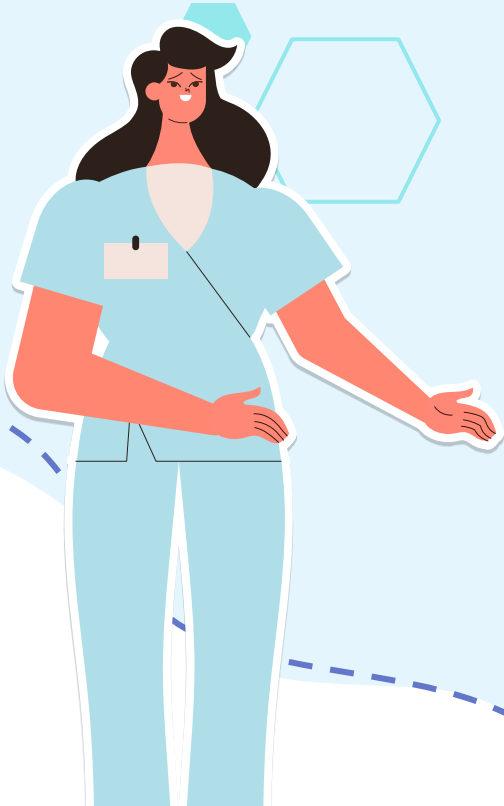
Mars

Despite being red, Mars is actually a cold place. It's full of iron oxide dust

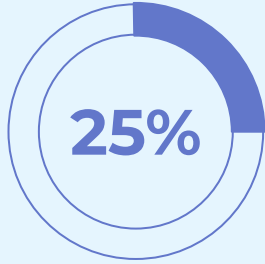
Patient satisfaction

95%

Mercury is the closest planet to the Sun and the smallest one in the Solar System—it's only a bit larger than the Moon



Innovations



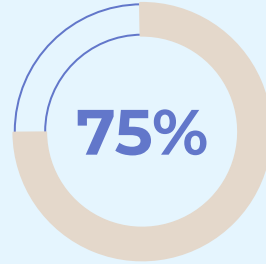
Mercury

It's the closest planet to the Sun and the smallest one in the Solar System



Venus

Venus has a beautiful name and is the second planet from the Sun



Mars

Despite being red, Mars is actually a cold place. It's full of iron oxide dust

Quality improvement measures



Mercury

It's the closest planet to the Sun



Venus

Venus is the second planet from the Sun



Mars

Mars is actually a very cold place



Jupiter

Jupiter is the biggest planet of them all



Saturn

It's composed of hydrogen and helium



Neptune

It's the farthest planet from the Sun

Testimonials

Jenna Doe, 25

"Neptune is the fourth-largest planet by diameter in the entire Solar System"

John James, 36

"Mercury is the closest planet to the Sun and the smallest one in the Solar System"



Awards we have achieved

Category	Year	Description
Service	2018	Mercury is a very small planet
Innovation	2019	Venus has a beautiful name
Treatments	2020	Mars is actually a cold place

Awesome words





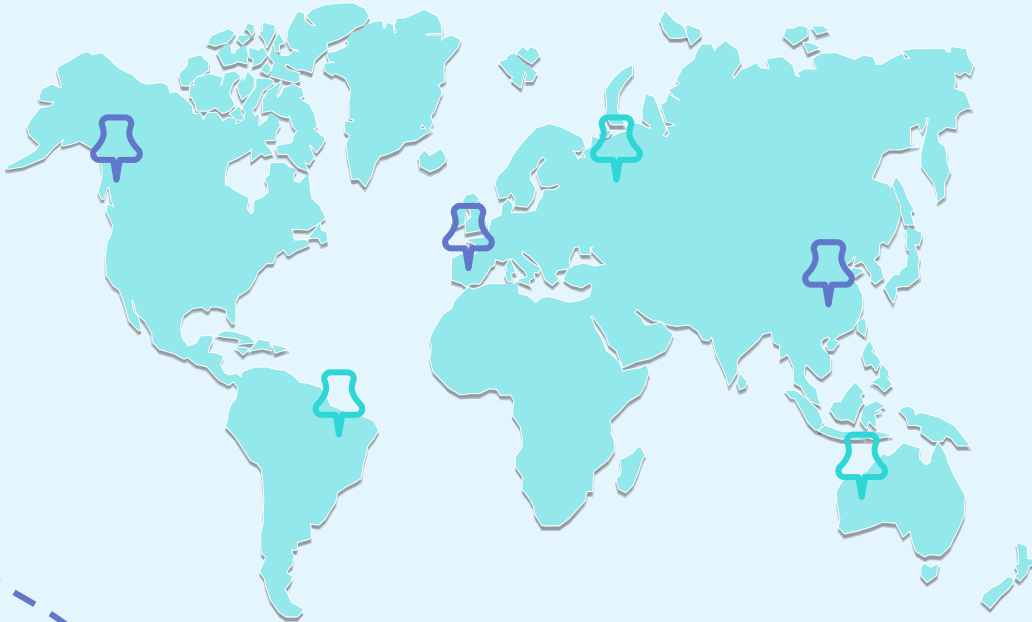
**A picture is worth a
thousand words**

A picture reinforces the concept

Images reveal large amounts of data, so
remember: use an image instead of a long text.
Your audience will appreciate it



Our centers



Current centers

Mars is actually a very cold place



Future centers

Venus is the second planet from the Sun

Desktop mockup

You can replace the image on the screen with your own work. Just right-click on it and select "Replace image"

