PROJECT AKHIR

PENGANTAR MULTIVARIAT



Analisis Statistik MANOVA dan ANOVA pada data Impor Bahan Baku dan Bahan Penolong tahun 1996-2022

Dosen Pembimbing:

Prof. Dra. Titin Siswantining, D.E.A.

Disusun oleh:

Vanny Khairunnisaa

2206051506

DEPARTEMEN MATEMATIKA FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM UNIVERSITAS INDONESIA

1. Abstrak

Impor bahan baku dan barang penolong memegang peranan penting bagi kelangsungan industri suatu negara. Menurut data impor dari Badan Pusat Statistik (BPS) mengenai bahan baku dan barang penolong dari tahun 1996 hingga 2022 menyoroti perubahan dalam tiga sektor utama: Makanan dan Minuman untuk Industri, Bahan Baku untuk Industri, serta Bahan Bakar dan Pelumas. Analisis statistik MANOVA dan ANOVA dari mata kuliah Pengantar Multivariat diharapkan memberikan wawasan mendalam mengenai impor bahan baku utama dan olahan dari ketiga sektor industri utama di Indonesia.

Analisis MANOVA menegaskan adanya perbedaan signifikan dalam impor bahan baku dan barang penolong di ketiga sektor industri. Uji ANOVA juga menunjukkan perbedaan yang signifikan antara jumlah bahan baku utama dan olahan yang diimpor selama periode 1996-2022. Informasi ini memberikan pemahaman mendalam mengenai dinamika impor bahan baku dan penolong, pentingnya sektor-sektor ini bagi ekonomi Indonesia, serta implikasinya dalam pengambilan keputusan kebijakan industri.

2. Pendahuluan

2.1. Latar Belakang

Dalam bidang ekonomi dan perdagangan dunia faktor impor bahan baku dan barang penolong memiliki peranan penting dalam suatu negara dalam menjaga kelancaran dan kelangsungan industri dalam negeri. Nilai impor mempengaruhi produktivitas dan daya saing industri suatu negara dan dampaknya terhadap ekonomi global.

Melalui data "Impor Bahan Baku dan Bahan Penolong, 1996-2022" yang diterbitkan oleh Badan Pusat Statistik (BPS), mencakup tiga industri utama yaitu Makanan dan Minuman untuk Industri, Bahan Baku untuk Industri, serta Bahan Bakar dan Pelumas. Didapatkan data bahan baku dan bahan penolong utama serta olahan berdasarkan periode tahun 1996-2022 yang berubah.

Penggunaan analisis statistik dari mata kuliah Pengantar Multivariat seperti MANOVA dan ANOVA, diharapkan dapat memberikan informasi lebih dari data impor bahan baku utama dan olahan dari ketiga industri utama Indonesia

2.2. Rumusan Masalah

- Apakah terdapat perbedaan rata-rata yang signifikan terhadap jumlah impor barang antar industri Makanan dan Minuman, Bahan Baku, serta Bahan Bakar dan Pelumas pada periode tahun 1996-2022 di Indonesia?
- 2) Apakah terdapat perbedaan rata-rata jumlah impor barang bahan baku utama dan olahan pada ketiga industri utama yang signifikan pada periode tahun 1996-2022 di Indonesia?

2.3. Tujuan

- Mengetahui apakah ada perbedaan rata-rata yang signifikan jumlah impor barang antar industri pada periode 1996-2022 di Indonesia menggunakan uji MANOVA;
- Mengetahui apakah terdapat perbedaan rata-rata yang signifikan antara jumlah impor bahan baku utama dan olahan pada periode 1996-2022 di Indonesia menggunakan uji ANOVA; dan
- 3) Mencari informasi yang terdapat pada data.

3. Metode Penelitian

3.1. Pengertian MANOVA

MANOVA (*Multivariate Analysis of Variance Model*) merupakan salah satu metode analisis data statistik yang dipelajari pada mata kuliah Pengantar Multivariat yang merupakan perluasan dari ANOVA. Untuk melakukan uji MANOVA dibutuhkan lebih dari satu variabel dependen untuk mengetahui apakah terdapat perbedaan pengaruh dari variabel dependen tersebut. Uji ini dilakukan jika peneliti menginginkan satu kali uji pada sekumpulan variabel dan mengetahui bagaimana variabel independen dapat mempengaruhi pola variabel dependennya.

3.1.1. Asumsi

Asumsi yang harus dipenuhi pada uji MANOVA adalah sebagai berikut:

- 1) Data berdistribusi normal multivariat
- 2) Homogenitas matriks varians kovarians atau homoskedastisitas
- 3) Terdapat korelasi antar variabel atau variabel saling bebas (independent)

3.1.2. Hipotesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

 H_1 : setidaknya terdapat dua μ yang berbeda

3.1.3. Statistik Uji

Pada uji MANOVA terdapat beberapa statistik uji yang dapat digunakan dalam membuat keputusan. Sebelum menggunakan statistik uji, akan dicari matriks "between" dan "within" atau matriks **H** dan **E** dengan rumus berikut:

• SSH =
$$n \sum_{i=1}^{k} (\overline{y}_{i.} - \overline{y}_{..})^2 = \sum_{i=1}^{k} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{kn}$$

• SSE =
$$\sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \overline{y}_{i})^2 = \sum_{ij} y_{ij}^2 - \sum_{i=1}^{k} \frac{y_{i}^2}{n}$$

•
$$\mathbf{H} = n \sum_{i=1}^{k} (\overline{y}_{i.} - \overline{y}_{..}) (\overline{y}_{i.} - \overline{y}_{..})' = \sum_{i=1}^{k} \frac{1}{n} y_{i.} y_{i.}' - \frac{1}{kn} y_{..} y_{..}'$$

•
$$\mathbf{E} = \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \overline{y}_{i.})(y_{ij} - \overline{y}_{i.}) = \sum_{ij} y_{ij} y_{ij} - \sum_{i=1}^{k} \frac{1}{n} y_{i.} y_{i.}$$

Berikut beberapa statistik uji dalam uji MANOVA:

1) Wilks' Test

Statistik uji

$$|\Lambda| = \frac{|E|}{|E+H|}$$

Aturan Penolakan

 H_0 ditolak jika $\Lambda \leq \Lambda_{\alpha:v:VH:VE}$ atau $p-value < \alpha$

<u>Keterangan</u>

$$V_{H} = k - 1, V_{E} = k(n - 1)$$

2) Roy's Test

Statistik uji

$$\theta = \frac{\lambda 1}{1 + \lambda 1}$$

Aturan Penolakan

$$H_0$$
 ditolak jika $\theta \ge \theta_{\alpha:s:m:N}$ atau $p - value < \alpha$

Keterangan

$$V_{H} = k - 1, V_{E} = k(n - 1), s = min(p, V_{H}), m = \frac{|VH - P| - 1}{2}, N = \frac{VE - P - 1}{2}$$

3) Pillai Test

Statistik uji

$$V^{(s)} = tr[(E + H)^{-1}H]$$

Aturan Penolakan

$$H_0$$
 ditolak jika $V^{(s)} \ge V_{\alpha}^{(s)}$ atau $p - value < \alpha$

<u>Keterangan</u>

$$\overline{V_H = k - 1}, s = min(p, V_H)$$

4) Hottelling Lawley

Statistik uii

$$U^{(s)} = tr[E^{-1}H]$$

Aturan Penolakan

$$H_0$$
 ditolak jika $\frac{\mathit{VE}}{\mathit{VH}}\mathit{U}^{(s)}hitung \geq \frac{\mathit{VE}}{\mathit{VH}}\mathit{U}^{(s)}$ tabel atau $p-\mathit{value} < \alpha$

Keterangan

$$V_{H} = k - 1, V_{E} = k(n - 1), s = min(p, V_{H})$$

3.2. Pengertian ANOVA

ANOVA (*Univariate Analysis of Variance Model*) merupakan salah satu metode analisis data statistik yang dipelajari pada mata kuliah Pengantar Multivariat yang digunakan untuk menganalisis perbedaan rata-rata variabel independen antar satu variabel dependen. Model pada tiap observasi diberikan secara matematis sebagai $y_{ij} = \mu_i + \varepsilon_{ij}$

3.2.1. Asumsi

Asumsi berikut diperlukan untuk mendapatkan nilai statistik uji F

- 1) Data berdistribusi normal
- 2) Homogenitas matriks varians kovarians atau homoskedastisitas
- 3) Terdapat korelasi antar variabel atau variabel saling bebas (*independent*)
- 3.2.2. Hipotesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

 H_1 : setidaknya terdapat dua μ yang berbeda

3.2.3. Statistik Uji

$$F = \frac{ns_{y}^{2}}{s_{e}^{2}} = \frac{(\sum_{i=1}^{2} y_{i}^{2}/n - y_{j}^{2}/kn)/(k-1)}{(\sum_{ij} y_{ij}^{2} - \sum_{i} y_{i}^{2}/n)/[k(n-1)]} = \frac{SSH/(k-1)}{SSE/[k(n-1)]} = \frac{MSH}{MSE}$$

3.2.4. Aturan Keputusan

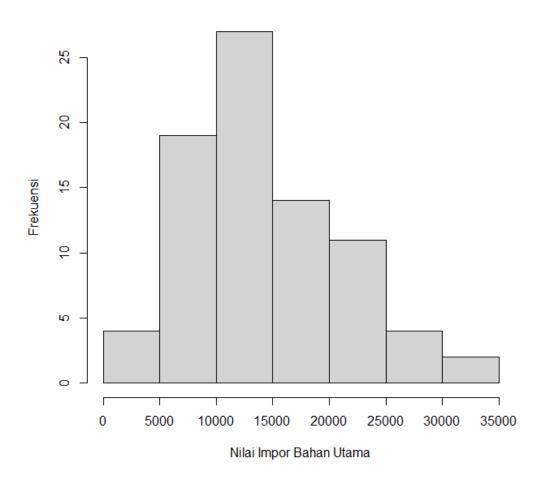
$$H_0$$
 ditolak jika $F > F_{\alpha;k-1;k(n-1)}$ atau $p - value < \alpha$

4. Pembahasan dan Hasil

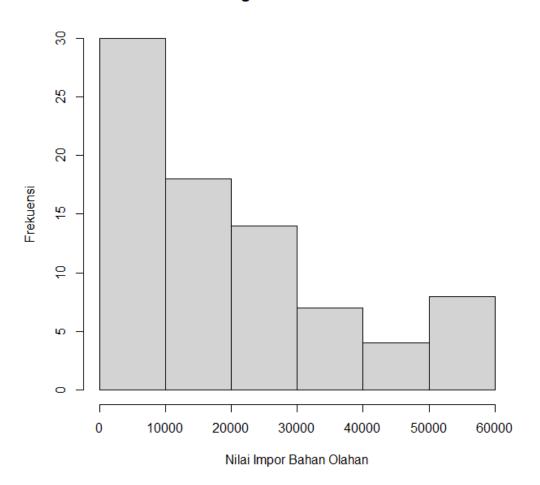
4.1. Statistik Deskriptif

Berikut adalah statistik deskriptif dari data numerik yang berisikan nilai minimum, median, mean, kuartil, dan maksimum dari kolom numerik (Utama dan Olahan). Data yang digunakan terdiri atas 3 kolom dan 81 baris. Untuk melengkapi statistik deskriptif, dibuat juga histogram dari kolom Utama dan kolom Olahan berikut

Histogram Bahan Utama







4.2. Uji Asumsi

4.2.1. Data berdistribusi normal (Normalitas)

Akan dilakukan uji asumsi normalitas pada data menggunakan shapiro test terhadap residual dan kedua variabel independen.

Hipotesis

 H_0 : Data berdistribusi normal

 H_1 : Data tidak berdistribusi normal

Taraf Signifikansi

 $\alpha = 0.05$

<u>Hasil Uji</u>

```
shapiro.test(res)
Shapiro-Wilk normality test
data: res
W = 0.99006, p-value = 0.0793
shapiro.test(df$Utama)
Shapiro-Wilk normality test
data: df$Utama
W = 0.95462, p-value = 0.05977
shapiro.test(df$olahan)
Shapiro-Wilk normality test
data: df$0lahan
W = 0.89108, p-value = 0.06466
```

Aturan Penolakan

 H_0 ditolak jika $p - value < 0.05 = \alpha$

Kesimpulan

Dari ketiga data yang diuji, didapatkan p - value > 0.05 yang berarti seluruh H_0 gagal untuk ditolak. Sehingga didapatkan asumsi normalitas terpenuhi.

4.2.2. Homogenitas matriks varians kovarians (Homoskedastis)

Akan dilakukan uji asumsi homoskedastis pada data menggunakan boxplot dan bartlett test terhadap residual dan kedua variabel independen.

Hipotesis

 H_0 : Data memiliki variansi yang konsisten (homoskedastis)

 H_1 : Data tidak tidak memiliki variansi yang konsisten (heteroskedastis)

Taraf Signifikansi

```
\alpha = 0.05
Hasil Uji
```

```
boxM(res, df$Industri)
Box's M-test for Homogeneity of Covariance Matrices
data: res
Chi-sq (approx.) = 1.7843, df = 2, p-value = 0.41
```

```
bartlett.test(df$Utama~df$Industri)

Bartlett test of homogeneity of variances

data: df$Utama by df$Industri
Bartlett's K-squared = 10.79, df = 2, p-value = 0.426

bartlett.test(df$Olahan~df$Industri)

Bartlett test of homogeneity of variances

data: df$Olahan by df$Industri
Bartlett's K-squared = 68.33, df = 2, p-value = 0.453
```

Aturan Penolakan

 H_0 ditolak jika $p - value < 0.05 = \alpha$

Kesimpulan

Dari ketiga tes yang dilakukan, didapatkan p-value>0.05 yang berarti seluruh H_0 gagal untuk ditolak. Sehingga didapatkan variansi data konsisten secara multivariat dan univariat atau asumsi homoskedastis terpenuhi.

4.2.3. Independensi

Akan dilakukan uji asumsi independensi pada data menggunakan durbin watson test terhadap residual data.

Hipotesis

 H_0 : Terdapat korelasi antar variabel (*independent*)

H₁: Tidak terdapat korelasi antar variabel (*dependent*)

Taraf Signifikansi

```
\alpha = 0.05
<u>Hasil Uji</u>
```

```
dwtest(res~fitted(model))
Durbin-Watson test

data: res ~ fitted(model)
DW = 0.33485, p-value = 0.9881
alternative hypothesis: true autocorrelation is greater than 0
```

Aturan Penolakan

 H_0 ditolak jika $p - value < 0.05 = \alpha$ Kesimpulan Dari tes yang dilakukan, didapatkan p-value>0.05 yang berarti H_0 gagal untuk ditolak. Sehingga didapatkan korelasi antar variabel atau asumsi independensi terpenuhi.

4.3. Uji MANOVA

4.3.1. Hipotesis

 H_0 : $\mu_1 = \mu_2 = \mu_3$ (tidak terdapat beda mean signifikan antar industri)

 H_1 : setidaknya ada dua beda μ (ada beda mean signifikan antar industri)

4.3.2. Taraf Signifikansi

$$\alpha = 0.05$$

- 4.3.3. Statistik Uji
 - 4.3.3.1. Mean μ_1 dan μ_2

Dengan bantuan program R, didapatkan μ_1 , μ_2 , dan μ_3 sebagai

Bahan Bakar dan Pelumas Bahan Baku Industri Makanan dan minuman y1 16727.36 17185.12 8461.926 y2 20896.13 34780.46 3732.737

4.3.3.2. Matriks H dan matriks E

Dengan bantuan program R, didapatkan matriks H dan E sebagai berikut

4.3.3.3. Wilks' Test

<u>Uji MANOVA dengan Wilks' Test menggunakan R</u>

Nilai Wilks' stat menggunakan R

```
> Lambda <- det(E)/det(E + H); Lambda
           [1] 0.2917086
          Statistik uji
          |\Lambda| = \frac{|E|}{|E+H|} = 0.2917086
          Aturan Penolakan
                ditolak jika \Lambda = 0.2917 \le 0.888 \approx \Lambda_{0.05:2:2:78}
          H_{\circ}
                                                                       atau
          p - value < 0.05 = \alpha
          Keterangan
          V_{H} = 3 - 1 = 2, V_{E} = 3(27 - 1) = 78
4.3.3.4.
          Roys' Test
          Uji MANOVA dengan Roys' Test menggunakan R
           > summary(manova, test ="Roy")
                            Roy approx F num Df den Df Pr(>F)
           Industri 2 1.8109 70.626 2 78 < 2.2e-16 ***
           Residuals 78
           Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
          Nilai Roy's' stat menggunakan R
           > EinvH.eigen <- eigen(solve(E) %*% H)
           > roy.stat <- EinvH.eigen$values[1]; roy.stat
           [1] 1.810929
           > theta <- roy.stat/(1+roy.stat); theta
           [1] 0.6442458
          Statistik uji
          \theta = \frac{\lambda 1}{1 + \lambda 11} = 0.6442458
          Aturan Penolakan
              ditolak jika \theta = 0.644 \ge 0.122 \approx \theta_{0.05:2:0:37.5}
          H_{\circ}
                                                                       atau
          p - value < \alpha
          Keterangan
          V_{H} = 3 - 1 = 2, V_{E} = 3(27 - 1) = 78,
          s = min(p, V_H) = 2, m = \frac{|2-2|-1}{2} = 0, N = \frac{78-2-1}{2} = 37.5
4.3.3.5.
          Pillai Test
          Uji MANOVA dengan Pillais' Test menggunakan R
           > summary(manova, test ="Pillai")
                       Df Pillai approx F num Df den Df Pr(>F)
           Industri 2 0.82427 27.342 4 156 < 2.2e-16 ***
           Residuals 78
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Nilai Pillais' stat menggunakan R

Statistik uji

$$V^{(s)} = tr[(E + H)^{-1}H] = 0.8242734$$

Aturan Penolakan

$$H_0$$
 ditolak jika $V^{(2)}=0.8243 \ge 0.218 \approx V_{0.05}^{(2)}$ atau $p-value < \alpha$

<u>Keterangan</u>

$$V_{H} = 3 - 1 = 2$$
, $s = min(p, V_{H}) = 2$

4.3.3.6. Hotelling-Lawley Test

<u>Uji MANOVA dengan Hotelling-Lawley Test menggunakan R</u>

Nilai Hotelling-Lawley stat menggunakan R

Statistik uji

$$U^{(s)} = tr[E^{-1}H] = 2.0304$$

Aturan Penolakan

$$H_0$$
 ditolak jika $\frac{VE}{VH}U^{(2)}hitung = 79.186 \ge 5.02 \approx \frac{VE}{VH}U^{(2)}tabel$ atau $p-value < \alpha$

<u>Keterangan</u>

$$V_H = 3 - 1 = 2$$
, $s = min(p, V_H) = 2$, $V_E = 3(27 - 1) = 78$

4.3.4. Hasil Uji

Didapatkan pada ke-4 statistik uji MANOVA memenuhi aturan penolakan dimana H_0 ditolak. Maka didapatkan bahwa setidaknya terdapat perbedaan mean pada jumlah impor industri utama Indonesia periode 1996-2022.

4.4. Uji ANOVA

4.4.1. Hipotesis

$$H_0: \mu_1 = \mu_2$$

(tidak ada beda mean signifikan antara bahan utama dan bahan olahan)

$$H_1: \mu_1 \neq \mu_2$$

(ada beda mean signifikan antara bahan utama dan bahan olahan)

4.4.2. Taraf Signifikansi

$$\alpha = 0.05$$

4.4.3. Statistik Uji

Akan digunakan uji F sebagai statistik uji. Dengan menggunakan bantuan program R, didapatkan hasil uji ANOVA seperti berikut

4.4.4. Aturan Keputusan

```
H_0 ditolak jika F > F_{\alpha;k-1;k(n-1)} atau p-value < 0.05 = \alpha Menggunakan R, didapatkan nilai F_{0.05;2;78} = 3.113792 > qf (0.95, 2, 78) [1] 3.113792
```

Maka dengan $F=21.63>3.113=F_{0.05:2:78}$ miliki bahan baku utama dan $F=57.65>3.113=F_{0.05:2:78}$ miliki bahan baku olahan. Kedua H_0 ditolak sehingga H_1 diterima atau terdapat beda mean signifikan antara bahan utama dan bahan olahan.

5. Simpulan

- Nilai impor bahan baku utama terkecil berada pada nilai 3.909 ton pada industri Makanan dan Minuman di tahun 1998 dengan nilai impor tertinggi sebesar 32.790 ton pada industri Bahan Baku Industri di tahun 2022. Nilai impor bahan baku utama pada periode 1996-2022 memiliki rata-rata 14.125 ton.
- Nilai impor bahan baku olahan terkecil berada pada nilai 906,6 ton pada industri
 Makanan dan Minuman di tahun 1996 dengan nilai impor tertinggi sebesar

- 57.069,8 ton pada industri Bahan Baku Industri di tahun 2022. Nilai impor bahan baku olahan pada periode 1996-2022 memiliki rata-rata 19.803,1 ton.
- Dengan menggunakan uji MANOVA, didapatkan bahwa ketiga industri (Makanan dan Minuman, Bahan Baku Industri, serta Bahan Bakar dan Pelumas) impor bahan baku dan bahan penolong pada periode 1996-2022 memiliki beda mean yang signifikan.
- Dengan menggunakan uji ANOVA, didapatkan bahwa terdapat beda mean yang signifikan antara jumlah bahan baku utama dengan bahan baku olahan pada impor bahan baku dan bahan penolong periode 1996-2022.

Lampiran 1

Berikut adalah data yang digunakan dalam makalah ini yang bersumber dari Badan Pusat Statistik (BPS) <u>Badan Pusat Statistik (bps.go.id)</u>

Impor Bahan Baku dan Barang Penolong dalam ton, 1996-2022

Tahun	Jumlah	Makanan dan Minuman untuk Industri		Bahan Baku untuk Industri		Bahan Bakar dan Pelumas	
		Utama	Olahan	Utama	Olahan	Utama	Olahan
1996	53.344,3	5.220,9	906,6	9.405,5	17.848,5	9.762,2	9.134,2
1997	55.410,5	4.750,7	1.059,4	9.243,4	18.854,2	9.513,8	10.932,2
1998	46.455,7	3.909,4	1.176,6	8.001,1	11.919,3	10.533,7	10.416,3
1999	54.175,4	4.973,6	1.958,7	10.593,2	13.906,0	11.695,8	10.606,1
2000	61.076,8	5.310,4	1.954,8	11.608,8	18.028,9	11.614,3	11.772,0
2001	60.245,1	4.139,4	1.747,7	12.220,5	17.319,2	14.206,0	9.843,9
2002	65.989,3	5.953,2	1.416,2	11.469,2	17.548,0	15.900,8	12.939,4
2003	63.952,7	5.080,1	1.724,7	10.865,5	16.691,5	20.240,5	8.637,7
2004	75.357,5	5.921,4	1.322,3	13.360,3	20.742,6	19.028,4	14.025,0
2005	76.582,9	5.780,0	2.105,4	10.588,2	22.548,8	15.748,5	18.665,0
2006	77.353,7	5.946,4	1.948,6	12.953,3	23.520,8	14.754,8	17.121,5
2007	81.741,0	6.467,2	3.113,0	12.363,9	26.016,6	15.222,2	17.495,9
2008	90.686,2	6.133,6	1.675,2	13.929,7	33.451,1	12.868,1	20.695,3
2009	84.720,0	6.413,1	2.950,1	11.915,9	27.571,8	15.381,7	19.182,8
2010	101.817,6	7.067,3	3.208,9	16.062,3	35.061,1	14.346,1	24.390,7
2011	116.101,8	8.207,2	4.282,5	18.811,0	41.188,5	13.316,2	28.313,5
2012	124.955,5	8.565,5	4.558,3	16.528,3	50.329,6	12.656,7	29.445,6
2013	132.395,7	9.026,4	4.935,2	20.431,0	48.510,5	16.534,6	30.741,0
2014	138.827,9	10.067,2	4.502,7	23.163,1	49.892,1	18.458,8	30.588,3
2015	139.139,6	10.183,6	4.818,6	22.644,0	50.431,7	21.540,8	27.564,7
2016	142.586,1	13.338,0	6.468,9	20.463,7	50.249,2	23.497,8	26.499,4
2017	151.882,7	15.206,6	6.400,9	23.673,3	51.735,2	22.302,2	30.496,2
2018	160.316,1	14.048,3	7.111,1	27.390,9	56.295,7	22.290,5	30.651,5
2019	153.275,3	14.771,7	6.581,7	27.381,2	56.090,4	18.526,3	27.669,7
2020	142.818,4	13.888,9	7.508,4	24.953,5	49.653,3	18.519,6	26.392,1
2021	167.768,5	15.123,1	7.432,4	31.187,9	56.597,9	27.215,2	27.761,3
2022	171.913,0	12.978,8	7.915,0	32.789,5	57.069,8	25.963,1	32.214,1