**Mobile Phones**

**Project Paper**

**Machine Learning**



| No | Nama Lengkap | NIM |
| --- | --- | --- |
| 1 | Rafael Adriel Wijaya | 2602162781 |
| 2 | Jacky Tan | 2602102265 |
| 3 | Reinhart Adriano Onggo | 2602162762 |
| Dosen | SYARIFAH DIANA PERMAI, S.Si,M.Si | D5309 |

**Automotive and Robotics Engineering**

**BINUS ASO School of Engineering**

**2024**

1. **Introduction**

In today's digital era, smartphones have become indispensable, transforming our methods of communication, work, and entertainment. Evolving from basic mobile phones with limited features, smartphones have developed into highly advanced devices integrating multiple technologies and capabilities, phones have come a long way from then until now. The current market is saturated with a variety of smartphones, each tailored to meet the diverse needs and preferences of users. There are many different types of smartphones available, starting from their operating systems, intended audiences, and distinctive features. Smartphones are primarily classified by their operating systems, with Android and iOS being the most prominent. Android, known for its open-source nature and extensive customization options, leads the global market with a variety of devices from numerous manufacturers. Conversely, iOS, developed by Apple Inc., is praised for its seamless integration with other Apple products and its emphasis on security and user experience.

The problem we encountered is mostly about the vast options of different specifications, features, and types of smartphones available in the market, making it difficult for customers to choose the device with the right specification for their professions and daily life activities. For this we are developing a machine learning model involving clustering on phone specs that will categorize the different types and features, including price range, RAM size, battery life, etc. into groups of feature ranges for easier understandment. Our expected output is to construct a model that helps with …

In summary, the variety in the smartphone market mirrors the diverse needs and preferences of consumers. Understanding the different types of smartphones not only aids consumers in making informed purchasing decisions but also provides insights into the technological advancements and trends shaping the future of mobile communication. This paper will provide a comprehensive overview of the various types of smartphones, highlighting their key features, target audiences, and the technological innovations driving their evolution.

Are you overwhelmed by the wide range of smartphone options available in the market today? Choosing the right smartphone that meets your customers' tech needs can be a daunting task. With the rapid advancements in technology, staying up-to-date with smartphone specifications is essential. In this article, we will delve into the world of smartphone specifications to help you make an informed decision on behalf of your customer.

From processors and RAM to camera quality and battery life, every smartphone specification plays a significant role in how well it performs. Understanding these specifications and how they affect the overall user experience is key to finding the perfect smartphone for your customers.

1. **Literatures Review**

An article titled “Usage of Principal Component Analysis (PCA)” by IBM. (n.d.). [1], highlights the importance of PCA in multivariate data analysis, emphasizing its ability to reduce dimensionality, extract important information, and visualize complex relationships between variables. It also discusses various aids and tools that can be used to enhance the interpretation of PCA results, such as scree plots, permutation tests, and correlation plots. PCA emphasizes feature engineering and feature extraction, making it useful for summarizing data and explaining topical effects by reducing data existing in multiple dimensions into 2-3 important dimensions.

Likas *et al*. [2] in the article of “The global k-means clustering algorithm” explains more about the global k-means algorithm which is an incremental approach to clustering that dynamically adds one cluster center at a time through a deterministic global search procedure consisting of N (with N being the size of the data set) executions of the k-means algorithm from suitable initial positions. The article also further discusses the proposal of modifications of the method. The basic idea underlying the proposed method is that an optimal solution for a clustering problem with M clusters can be obtained using a series of local searches.

Another paper titled “Using K-Means Clustering to Cluster Provinces in Indonesia” [3]. Contains a research conducted using K-means clustering in clustering data, in which the data is population density, school participation rate of 13-15, human development index, and open unemployment rate of a province in Indonesia consisting of 34 provinces. The paper uses K-Means Clustering to cluster Indonesian provinces based on their geographical, economic, and demographic characteristics. The authors aim to identify patterns and relationships among the provinces and to provide insights into the clustering results. The clustering results show that the provinces can be grouped into several clusters based on their characteristics. For example, the authors identify a cluster of provinces with high population density and GDP per capita, while another cluster consists of provinces with lower population density and GDP per capita. The paper demonstrates the effectiveness of K-Means Clustering in grouping Indonesian provinces based on their characteristics. The results provide valuable insights for policymakers and business leaders to better understand the provinces and to develop targeted strategies for economic development and social welfare.

1. **Data**

| Column Name | Data Type |
| --- | --- |
| # | Integer |
| Name | Object (String) |
| Brand | Object (String) |
| Model | Object (String) |
| Battery capacity (mAh) | Integer |
| Screen size (inches) | Float |
| Touch Screen | Object(Bool) |
| Resolution x | Integer |
| Resolution y | Integer |
| Processor | Integer |
| RAM (MB) | Integer |
| Rear camera | Float |
| Front camera | Float |
| Operating system | Object (String) |
| Wi-Fi | Object (Bool) |
| Bluetooth | Object (Bool) |
| GPS | Object (Bool) |
| Number of SIMs | Integer |
| 3G | Object (Bool) |
| 4G/LTE | Object (Bool) |
| Price | Integer (in INR) |

1. **End to End Machine Learning Steps**
   1. Data Visualization

First, the dataset was visualized in a table form to see all the features inside the dataset which can be seen on the number of variables present in the form of columns. In this dataset, there are 22 columns.

* 1. Data Preprocessing

After the data was visualized, the dataset was then checked for any duplication and missing value. After it was checked, the dataset was filtered where the columns with data type besides float or integer were removed so the dataset that’s going to be used has either float or integer data types for every column (except the first column named ‘#’). The dataset after this process will be called a dataframe in the upcoming process.

* 1. PCA

After the dataframe had been filtered from non-integer or float columns, the dataframe was run through a process called Principal Component Analysis or PCA for short where the dataframe’s columns will be analyzed to determine which columns to use for the modeling dataframe. This was done by making all of the PCA components based on the number of Eigenvalues created. The number of Eigenvalues created is minus one than the total number of columns in the dataframe input. In this case, there are 11 columns which means the number of Eigenvalues created is 10.

Then, check for the values that have e-0… in the back which indicates that the actual decimal values are 0,…. Which means that the values are less than zero. These particular values are going to be removed.

Next, make the loading score of all columns for each component. After all components are ready, for each column which in this case are the rows, find the highest value for each row.

After all highest values were found for all rows, choose the component where it has the most amount of the highest values found previously. These rows with the highest values are going to be used in the dataframe for the making process of the machine learning model.

* 1. Outliers filter (Outliers were deleted via csv file)

After the dataframe after the PCA was created, the data needed to be checked for any outliers. These outliers are data points different from the rest, the differences owned by these data points can cause errors due to the errors in measurement, bad data collection, or underestimating variables in doing so. These outliers can be detected by displaying the data in the form of a boxplot graph and find the UCL and LCL to remove the outliers in the file directly or directly find the UCL and LCL and then removing the outliers above and below them respectively.

* 1. Clustering

After the Outliers were removed, the dataframe can be used for the making process of the machine learning model, which in this case was clustering. Hence its name, this process aims in classifying data points into clusters based on the features present in the data points. These clusters will later have their own name or classes to represent all data points that are present inside their cluster. This process starts by making the distortion table to determine the best amount of clusters that need to be created. This can be done by finding the first node that has the closest difference to another node after it. Then, make the clustering plot using two variables from the dataframe. If the dataframe has more than two features in it, there can be several plots based on the combinations of variables used from the dataframe.

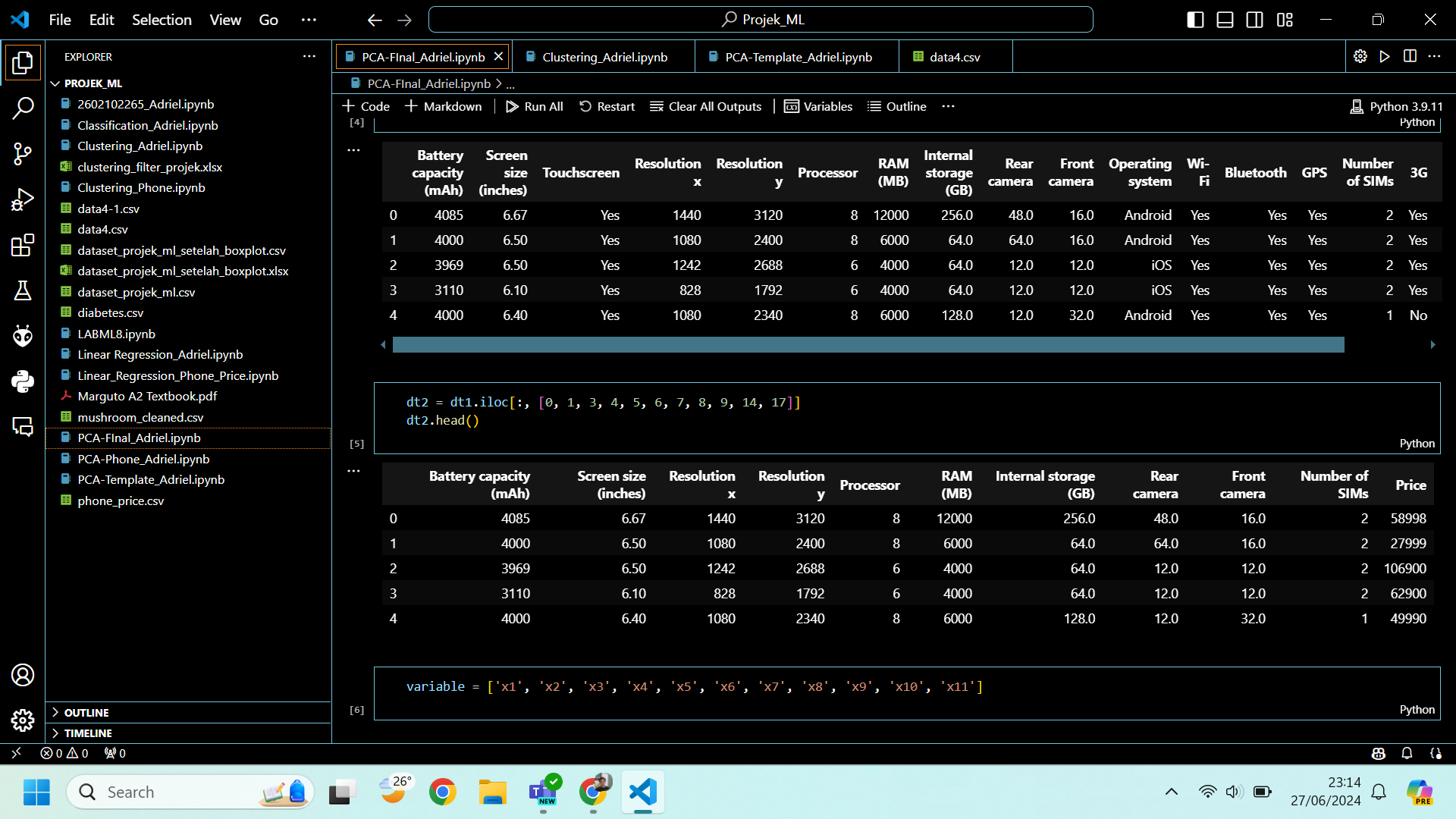
1. **Results**

Fig 1. Dataframe before PCA

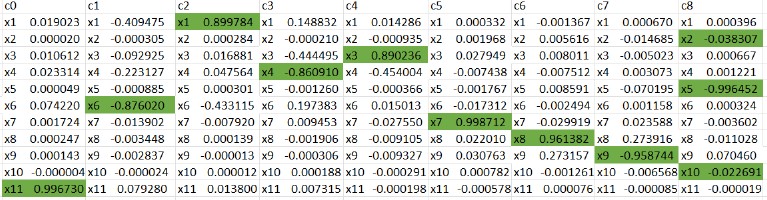


Fig 2. PCA Result

From the figure above, it can be concluded that component 8 (c8) will be used where the highest variables are x2, x5, and x10 which in contrast to the dataframe are Screen size (inch), Processor(number of processor), and Number of SIMs.

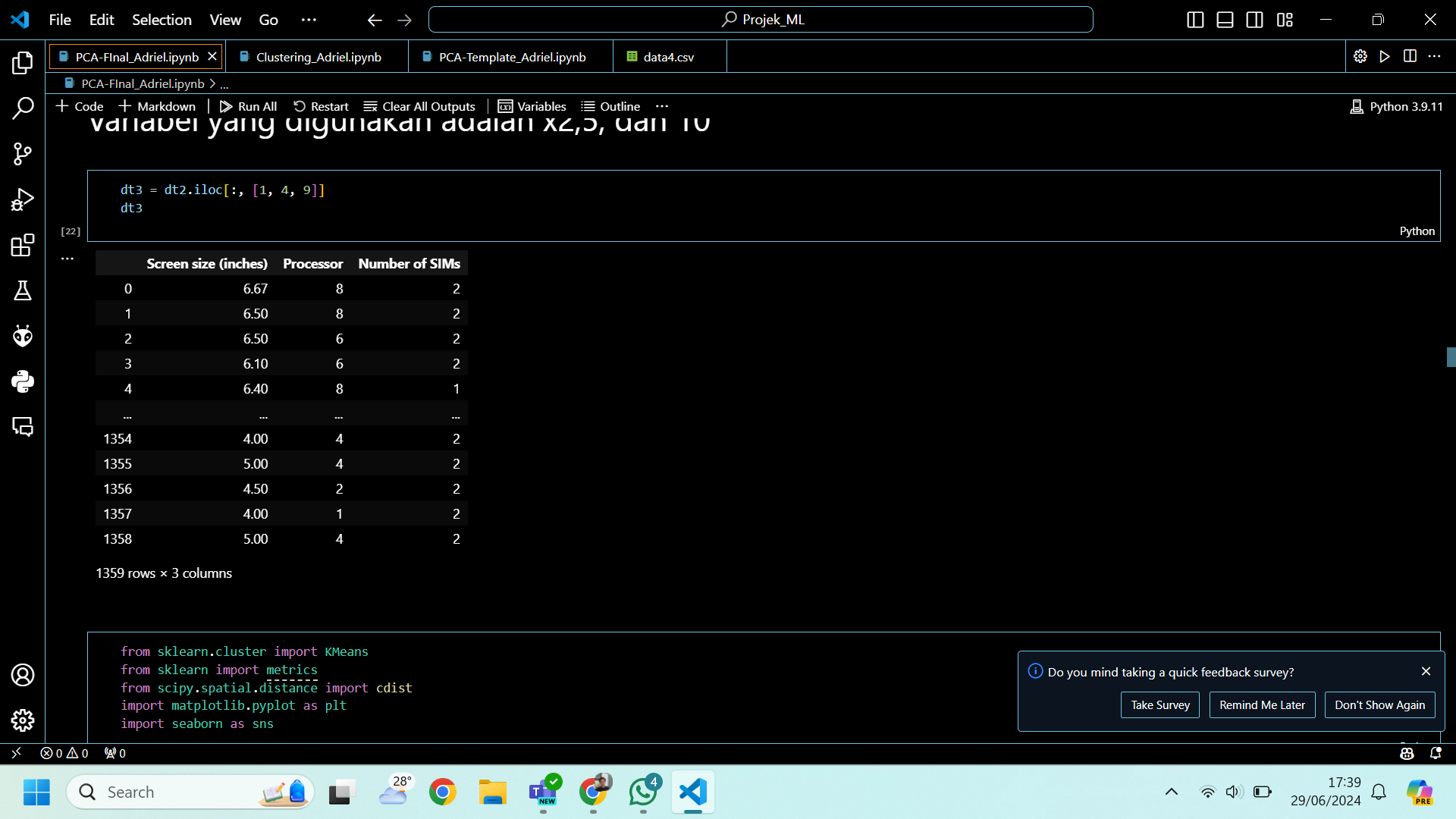
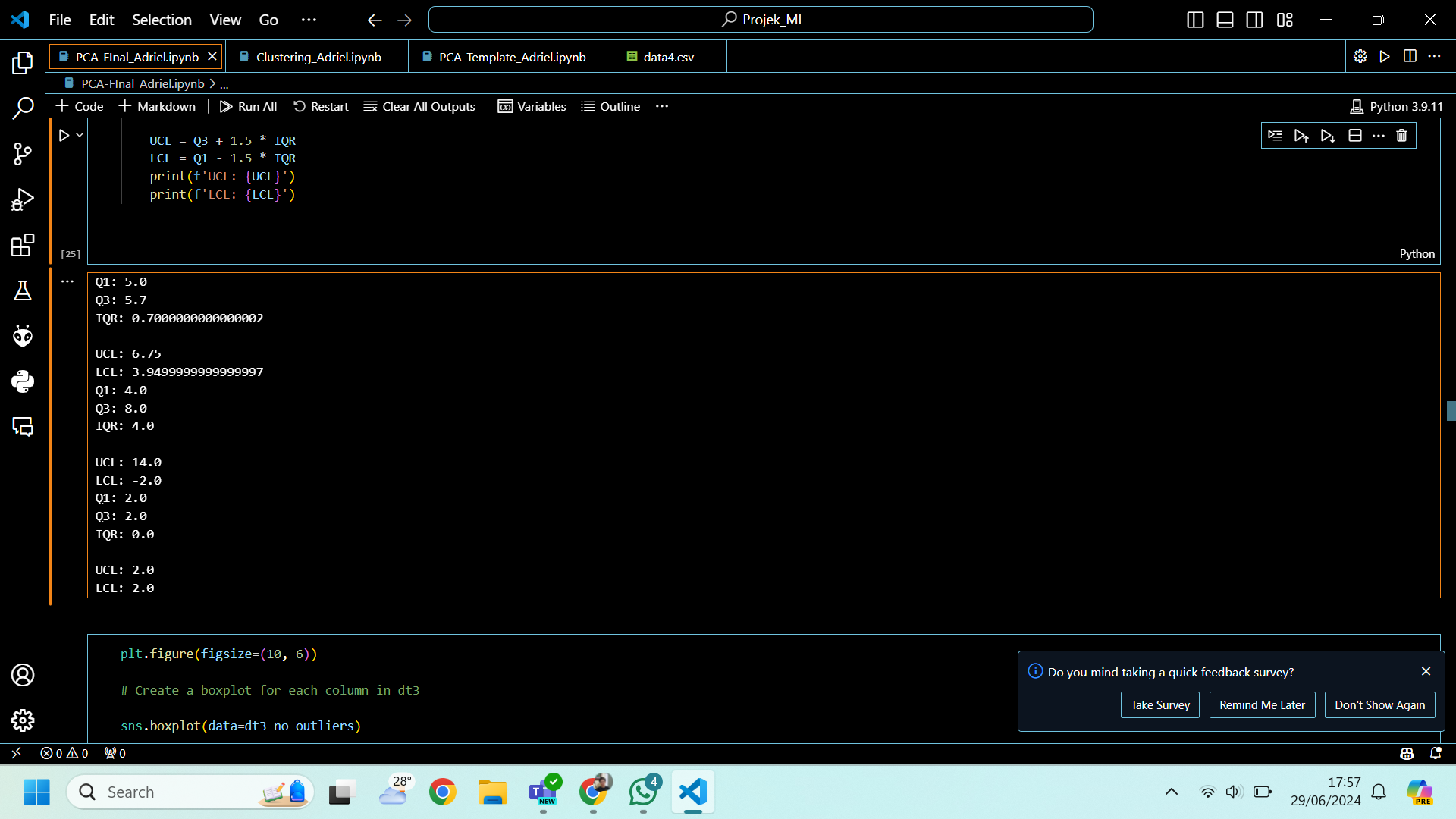


Fig 3. Dataframe after PCA



Fige 4. UCL and LCL for each column (from top to bottom: Screen size (inch), Processor(number of processor), and Number of SIMs)

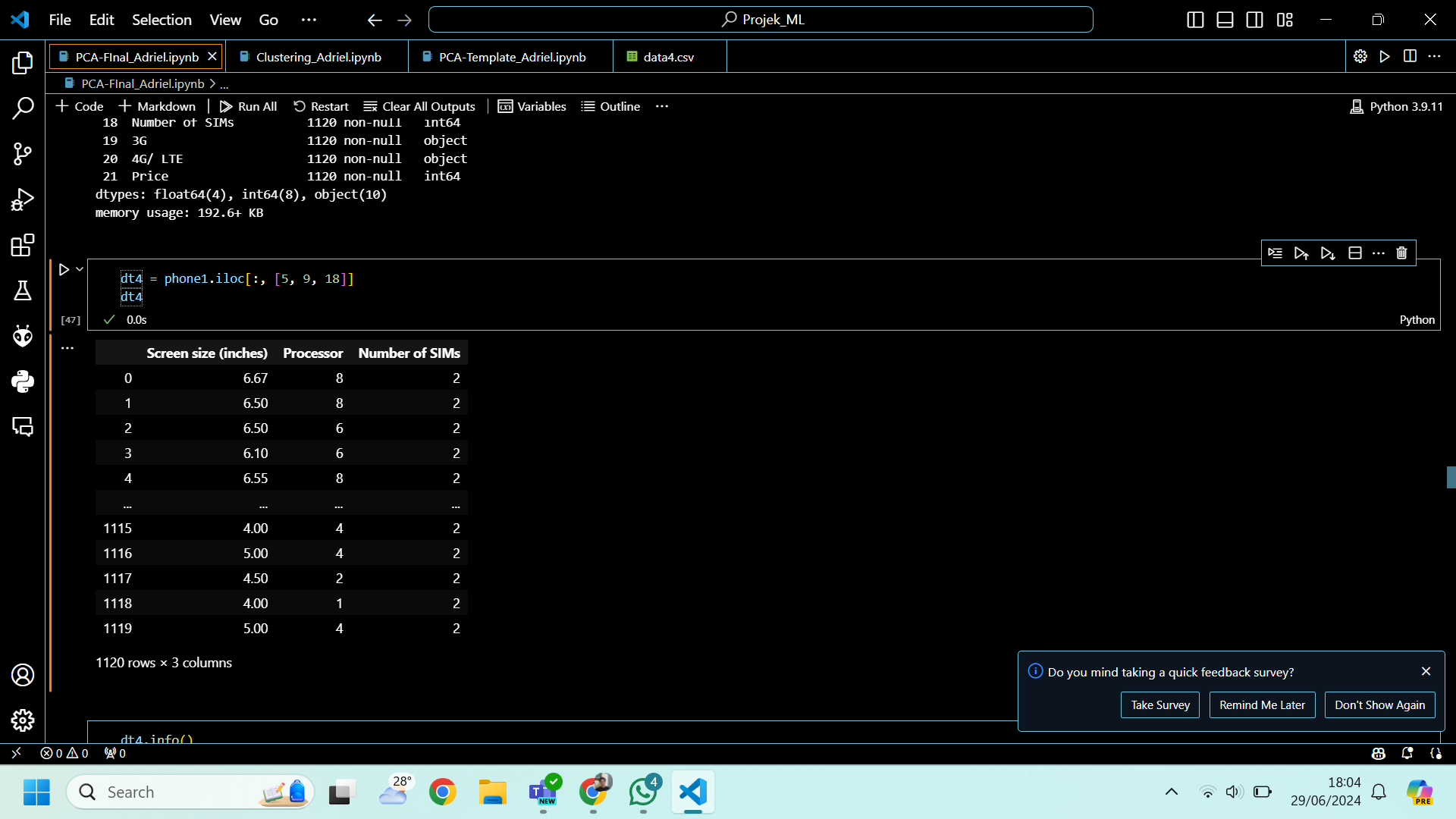


Fig 5. Dataframe after outliers were removed

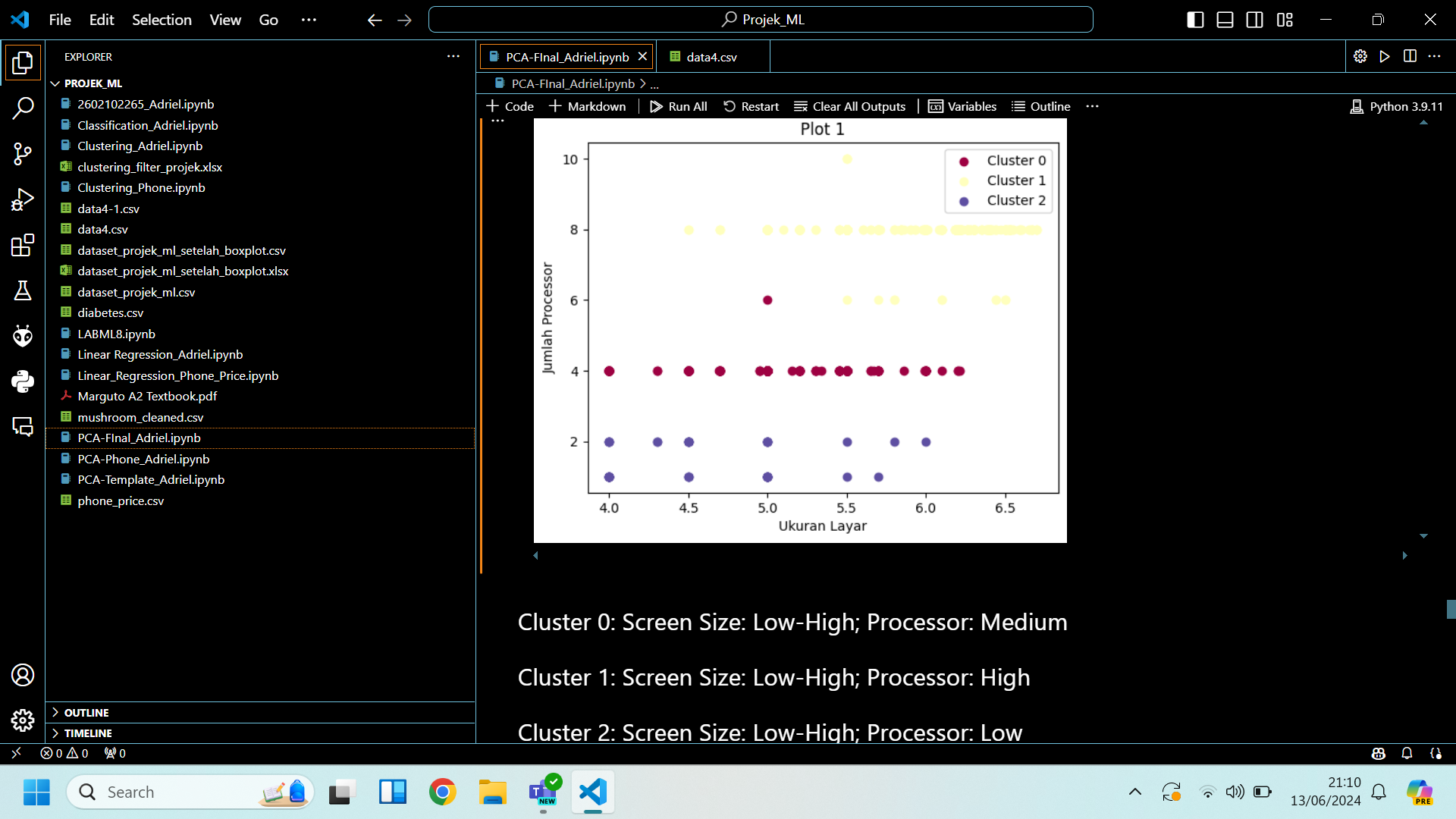


Fig 6. Plot 1 Clustering (Screen Size and Processor)

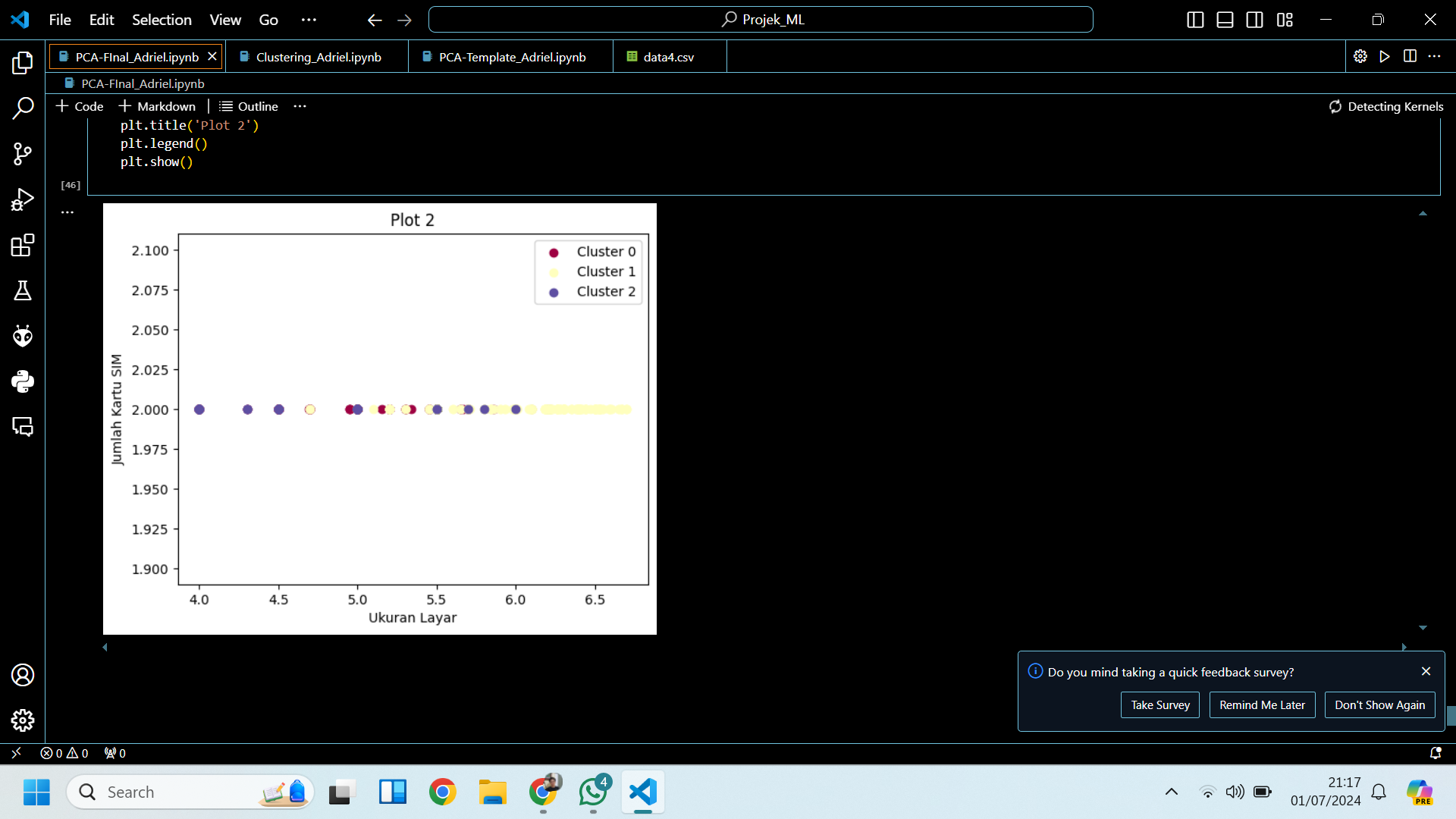


Fig 7. Plot 2 Clustering (Screen sizes and Number of SIMs)

From the figure above we can see that the clusters are overlapping each other because the “Number of SIMs” variable has only one type of integer which is ‘2’ and since there are no other integer beside it, whatever the value from the “Screen size” in the data points, some of them will overlap with each other whether they’re in the same clusters or not causing the clusters overlapping each other. This led to difficulties in finding the clusters along with the data points contained and its boundaries with other clusters. It also explains why plot 2 is not used.

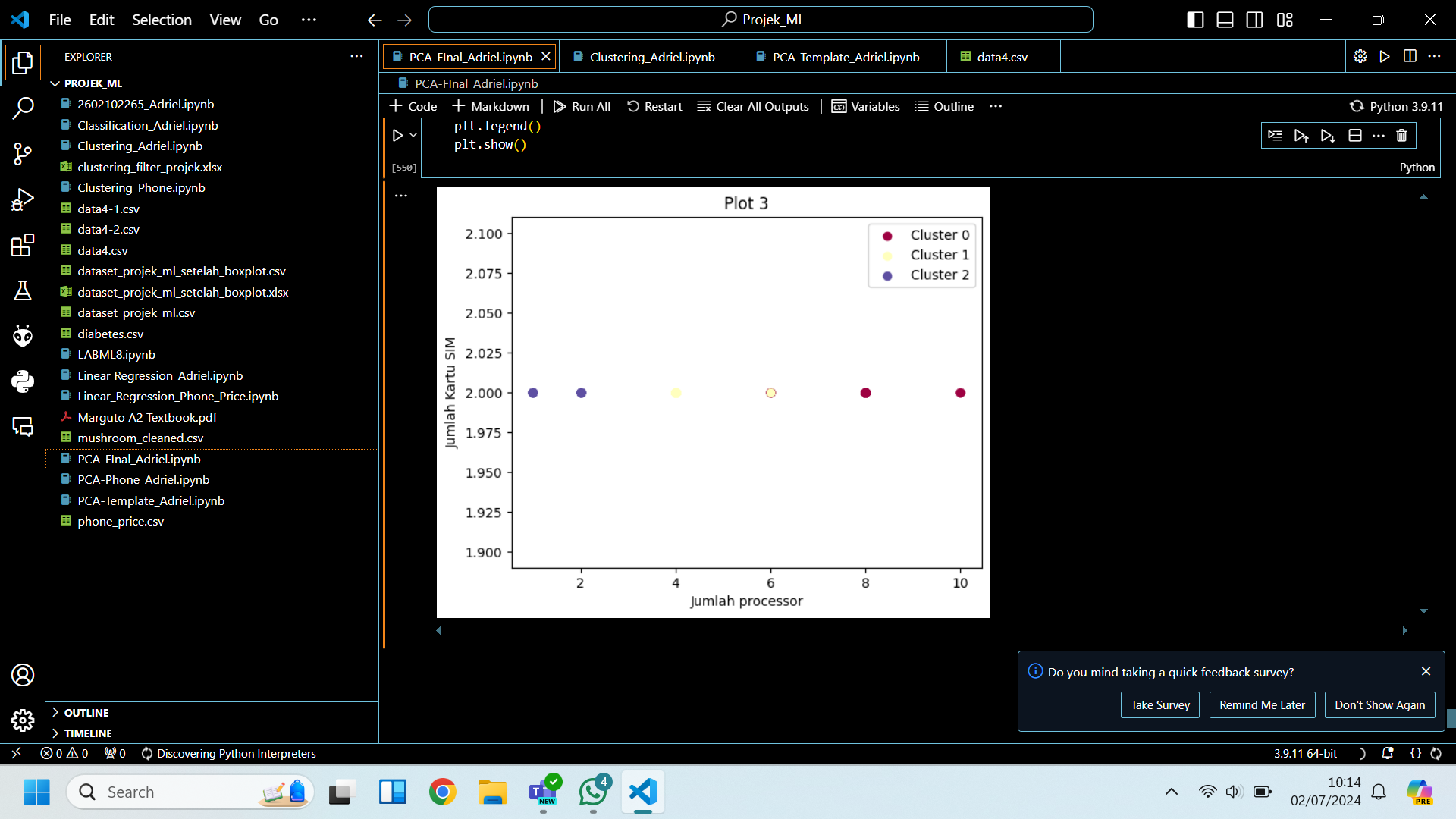


Fig 8. Plot 3 Clustering (Processor and Number of SIMs)

While in plot 3 on the figure above, despite the “Screen Size” variable being replaced by the “Processor” variable, there are still overlaps between two clusters, which occur between cluster 0 and 1(based on the plot in Fig 8). But unlike the one in plot 2 on Fig 7, the clusters are grouped and separated enough that they’re easier to identify their position and distances with other clusters despite two of them having some of their data points overlapped in the same position. Overall, plot 3 is also not used as well due the overlaps between cluster 0 and 1.

In conclusion, the plot that is going to be used is plot 1 because it's the only plot that has no overlaps between data points from other clusters so the clusters, along with the data points inside can be identified easily.

**Cluster Coordinate Plots**

|  | Cluster 0 | Cluster 1 | Cluster 2 |
| --- | --- | --- | --- |
| Screen Size | Low-High | Low-High | Low-High |
| Processor | Medium | High | Low |
|  |  |  |  |

Fig 9. Clustering result from Plot 1

| Cluster 0 | Cluster 1 | Cluster 2 |
| --- | --- | --- |
| 1. 10.or D/D2 2. Alcatel 3(X/V)/Pixi 4 (6)/U5 HD 3. Asus ZenFone:    1. 2 Deluxe/Laser(ZE(500/550)KL)    2. 3 Deluxe(ZS570KL)/Max(ZC520TL)    3. AR    4. Go (ZB450KL)/4.5(ZB452KG)/5.0 LTE/LTE(ZB500KL)/5.5 (ZB552KL)    5. Live    6. Max 4. Blu Life Mark/Win HD LTE 5. Cat S31 6. Celkon (Smart/Star)4G/UniQ 7. Comio C1/C1 Pro/C2/C2 Lite/P1/S1/S1 Lite/X1 Note 8. Coolpad Mega 2.5D/Note(3/5)Lite 9. Gionee Elife E3/F103/Marathon (M3/M4/M5 Lite)/P7/Pioneer P(3/3S/5 mini/5W/6)/S10 Lite/X1/X1s 10. Honor 4X/ Bee 2/Holly/Holly 2 Plus 11. HTC Desire (10 Lifestyle/310/500/600/630/816/816G/825)/U Ultra 12. Huawei Holly 4 Plus 13. iBall:     1. Andi (4.5 O'Buddy/5.5H Weber/5(G Blink 4G/N Dude/U Platino)/Avonte 5/HD6/Sprinter 4G)     2. Cobalt (5.5F Youva/Solus 4G) 14. Infinix Hot(4/6) Pro/Note 4/Smart (2/3 Plus) 15. InFocus Bingo (10/20 (M425)/21/50/50+)/M(260/370/535/812)/Turbo 5/Vision 3 16. Intex:     1. Aqua(3G Neo/4.0/4.5 Pro/4G (Mini/Strong)/5.5 (VR/VR+)/A(2/4)/Ace/Ace Mini/Amaze+/Classic/Costa/Craze/Craze II/Crystal/Crystal Plus/Dream II/Fish/Freedom/Glam/Jewel 2/Joy/Life III/Lions(4G/E3/N1/T1/X1+)/Lite/Music/Note 5.5/Power (4G/II/M)/Pride/Q(1+/7/7 Pro/7N/7N Pro/8)/Ring/S(1/2/3)/Secure/Sense (5.0/5.1)/Shine 4G/ Speed HD/Star 4G/Strong (5.1+/5.2)/Super/Supreme+/Trend/Trend Lite/Turbo 4G/Twist/View/Y2 (Power/Ultra)/Young/Zenith)     2. Cloud 4G(Smart/Star)/C1/Crystal 2.5D/Cube/Fame/Force/Gem+/Glory 4G/N/Pace/S9/String HD/Style 4G/Swift/Tread/V/Zest     3. Elyt Dual/E(6/7/1)     4. Infie (3/33)     5. Staari (10/11) 17. Itel A(20/22 Pro/44/44 Air/44 Power/44 Pro/45)/S(21/41/42)/Wish A(21/41/41+) 18. iVoomi Innelo/i(1/1s/2/2 Lite/Pro/V505)/Me(1/1+/2/3/3s/4/5)/V5/Z1 19. Jivi Revolution TnT3 20. Karbonn:     1. A((1/40) Indian/41 Power/91 Storm)     2. Aura (4G/Note (2/Play)/Power 4G Plus)     3. Fashion Eye     4. Frames S9     5. K9((Kavach/Music/Viraat)4G/Smart (Eco/Grand/Selfie/Yuva)     6. Machone Titanium S310     7. Quattro L52     8. Sparkle V     9. Titanium (Frames S7/Jumbo 2/Mach (Five/One Plus)/Moghul) 21. Kult 10/Ambition/Beyond/Gladiator/Impulse 22. Lava:     1. A(48/59/67/68/71/76+/77/79/82/88/97)     2. Flair Z1     3. Iris (504q/Atom/Atom 2X/Atom 3/Fuel (60/F1)/Pro 20/X1/X1 Selfie)     4. P7     5. Pixel (V1/V2)     6. V(2s/5)     7. X(10/11/17/28/3/38/41+/46/50+/81)     8. Z(10/40/41/50/53/70/80/81/90/91) 23. LeEco Le (Max 2/Pro 3 Elite) 24. Lenovo:     1. A(1000/2010/5/6000/6000 Plus/6000 Shot/6600/6600 Plus/7700)     2. P780     3. S(660/860)     4. Vibe (B/P1m/Z2 Pro)     5. Z2 Plus     6. Zuk Z1 25. Lephone W(15/7) 26. LG G((3/4) Stylus/5)/K(7i/8 (2017)/9)/L90 Dual/Max/V20/X Power 27. Lyf:     1. C(451/459)     2. F8     3. Flame (1/2/4/6/7/7s/8)     4. Wind (1/2/3/4/4S/5/6/7/7i/7S) 28. Meizu C9 29. MicroMax:     1. Bharat (2/2 Plus/2 Ultra/4/5/5 Infinity Edition/Go/)     2. Bolt (Q(326+/331/332/338/339)/S302/Selfie)     3. Canvas (1/2 (2017)/4/5 Lite/A1/Blaze (4G/4G+)/Doodle 4/Fire (4G/4G+/5)/Hue/Infinity/Juice (3/3+/4)/Mega (2/4G)/Pace 4G/Play 4G/Selfie (2/4)/Spark/Spark 3/Turbo/Turbo Mini/XP 4G/Xpress 4G)     4. Evok Power     5. Selfie 2     6. Spark (Go/Vdeo)     7. Unite (2/3/4 Plus)     8. Vdeo (1/2) 30. Microsoft Lumia 540 Dual SIM 31. Mobiistar CQ 32. Motorola Moto (E5/C/C Plus/E4/E4 Plus/G/G (Gen 2)/G (Gen 3)/G4 Play/Z) 33. M-tech TEZ4G 34. Nokia 1/2/2.3/3 35. Nubia (N1 Lite/Z11) 36. Nuu Mobile M3/Q(500/626) 37. OnePlus 3/3T 38. Oppo A37/Neo 7/R1 39. Panasonic:     1. Eluga (A(2/4)/Arc/Arc 2/I(2/2 Activ/3/3 Mega/4/5/9)/L2/Prim/Ray/Ray (500/530/550/X)     2. P(100/50 Idol/65 Flash/66 Mega/71/75/77/85/88/9/90/91/95)     3. T(30/33/44/44 Lite/45 4G/50) 40. Reach (Allure/Cogent/Opulent) 41. Samsung:     1. Galaxy (Grand 2/J(2 ((2017)/(2018)/Core/Pro)/3 Pro/4/5 ((2016)/Prime))/On(5/5 Pro/7/7 Pro)     2. Z(2/3/4) 42. Sansui Horizon 1 43. Sony Xperia (L2/XZ/XZs) 44. Spice (Dream Uno/Pinnacle FHD/Stellar 520) 45. Swipe:     1. Elite (2 Plus/3/4G/Note/Power/Pro/Sense/Star     2. Konnect (5.1/Grand/Neo 4G/Plus/Power/Star     3. Neo Power     4. Virtue 46. Tambo TA-3 47. TCL 560 48. Tecno:     1. Camon i(4/ACE/Ace (2/2X)/Sky 2/Twin)     2. i(3/3 Pro/5/5 Pro) 49. Videocon:     1. Cube 3     2. Delite 11+     3. Infinium Z(45 Amaze/51 (Nova+/Punch/Q Star))     4. Krypton (22/V50(DA/FG))     5. Metal Pro 2     6. Z55 Krypton 50. Vivo:     1. V1     2. Y(15/15S/(21/27/31/51)L/53/53i/55s/71 51. Xiaomi:     1. Mi 5     2. Redmi (1S/2/(2/Note) Prime/(4/6)A) 52. Xolo:     1. Era/Era (1X/1X Pro/2/2(V/X)/3X/4(K/X)/X)     2. One HD     3. Q(1000 Opus/1010i/1100/2000/2100/3000     4. Win Q900s 53. Yu:     1. Ace     2. Yunique/Yunique (2/Plus)     3. Yuphoria     4. Yureka S 54. Zen:     1. Admire (Joy/Metal/Star/Swadesh/SXY/Unity     2. Cinemax (2/2+/3/Click/Force) 55. Ziox:     1. Astra (Colors/Curve/Force/Titan) 4G/Star     2. Duopix R1     3. QUIQ Flash 4G 56. Zopo:     1. Color F(1/2/5)/M(4/5)     2. Color X 5.5     3. Hero 1 57. ZTE (Blade A910/V5) 58. Zuk Edge | 1. 10.or E/G/G2 2. Acer Liquid Z630s 3. Alcatel A5 LED 4. Apple: 5. iPhone:  * XR * XS * XS Max * 11 Pro * 11 * 11 Pro Max  1. Asus: 2. 6Z 3. ROG Phone 4. ZenFone:  * 2 Laser (ZE601KL) * ZenFone 3 (ZE552KL) * Zenfone 3 Laser * ZenFone 3 Max (ZC553KL) * ZenFone 3S Max * 4 Selfie DC (ZD553KL) * 4 Selfie Pro * 5Z (ZS520KL) * Lite L1 (ZA551KL) * Max M1 (ZB556KL) * Max M2 * Max Pro M1 * Max Pro M2 * Selfie * Zoom S  1. Billion Capture+ 2. Blackshark 2 3. BlackBerry: 4. KEY2 5. KEY2 LE 6. Cat: 7. S41 8. S60 9. Coolpad: 10. Cool:  * 1 Dual * 3 * Play6  1. Dazen X7 2. Max 3. Note:  * 3 * 3 plus * 3S * 5 * 6  1. Gionee: 2. A1 3. A1 Lite 4. A1 Plus 5. Elife S7 6. F205 7. M7 power 8. Marathon M5 Plus 9. P5W 10. P7 Max 11. S Plus 12. S11 Lite 13. S6 14. S6 Pro 15. S6s 16. Google: 17. Pixel 4 18. Pixel 4 XL 19. Homtom: 20. H1 21. H3 22. H5 23. Honor: 24. 10 25. 10 Lite 26. 20i 27. 4C 28. 5C 29. 5X 30. 6 Plus 31. 6X 32. 7A 33. 7C 34. 7X 35. 8 36. 8 Lite 37. 9i 38. 9N 39. 9X 40. Holly 4 41. Play 42. View 10 43. HTC: 44. Desire:  * 526G+ Dual SIM * 616 Dual SIM * 626 Dual SIM * 628 Dual SIM * 728 Dual SIM * 728G Dual SIM * 820 * 820s * 828 Dual SIM  1. One:  * E9s Dual SIM * X9  1. UPlay 2. U11 3. U11 EYEs 4. U11+ 5. U12 Life 6. U12+ 7. Wildfire X 8. Huawei: 9. Mate:  * 2- pro * 30 Pro * 9  1. Nova:  * 3 * 3i  1. 20 Lite 2. 20 Pro 3. 30 Pro 4. Y9 Prime 2019 5. iBall Andi 5k Panther 6. Infinix: 7. Hot:  * 7 * 7 Pro * 8 * S3 * S3X  1. Note 5 2. S4 3. S5 4. S5 Lite 5. Zero 5 6. Zero 5 Pro 7. InFocus: 8. Epic 1 9. M535+ 10. M680 11. Turbo 5 Plus 12. Vision 3 Pro 13. Intex: 14. Aqua:  * GenX * Octa * Power  1. Karbonn: 2. Quattro L50 HD 3. Titanium Mach Two S360 4. Titanium Octane Plus 5. Lava: 6. Iris X8 7. Z25 8. Z71 9. Z93 10. LeEco: 11. Le 1s Eco 12. Le 2 13. Lenovo: 14. A7000 15. A7000 Turbo 16. K3 Note 17. K5 Note 18. K6 Note 19. K6 Power 20. K8 21. K8 Note 22. K8 Plus 23. P2 24. Vibe:  * P1 * P1 Turbo * S1 * Shot * X2  1. LG: 2. G7 ThinQ 3. G7+ ThinQ 4. G8s ThinQ 5. Q6 6. V30+ 7. V40 ThinQ 8. W30 9. W30 Pro 10. Lyf: 11. Earth 1 12. Earth 2 13. F1 14. F1S 15. Water:  * 1 * 2 * 3 * 7 * 7S * 8 * 9 * 10  1. Meizu: 2. 16 3. M16th 4. M3 Note 5. M5 Note 6. M6 Note 7. M6T 8. Pro 6 Plus 9. Pro 7 10. Mi A1 11. Micromax: 12. Canvas:  * 4 Pus * 5 * Infinity pro * Knight * Nitro * Nitro 4G * Pulse 4G * Selfie * Xpress 2  1. Dual 5 2. Infinity  * N11 * N12  1. iOne Note 2. Mobiistar XQ Dual 3. Motorola: 4. Moto:  * E5 Plus * E6s * G5S * G5S Plus * G6 * G6 Play * G6 Plus * G Turbo Edition * G4 Plus * G5 * G5 Plus * G7 * G7 power * G8 Plus * M * X Play * X Style * X4 * Z Play * Z2 Play  1. One:  * Action * Macro * Vision  1. mPhone: 2. 7 Plus 3. 8 4. Nokia: 5. 3.1 Plus 6. 5 7. 5.1 Plus 8. 6 9. 6.1 10. 6.1 Plus 11. 7 Plus 12. 7.1 13. 7,2 14. 8 15. 8.1 16. Nubia: 17. M2 Lite 18. M2 Play 19. N1 20. N2 21. Red Magic 22. Red Magic 3 23. Red Magic 3S 24. Z11 Mini 25. Z11 Mini S 26. OnePlus: 27. 5 28. 5T 29. 6 30. 6T 31. 7 32. 7 Pro 33. 7T 34. 7T Pro 35. 7T Pro Mclaren Edition 36. Oppo:  * A1K, A5, A57, A5s, A71, A71 (2018), A83, A83 (2018), A9, A9 2020, F1, F1 Plus, F3, F3 Plus, F5, F7, F9 Pro, Find X, K3, R15 Pro, R17 Pro, Reno, Reno 10x Zoom  1. Panasonic: 2. Eluga:  * A3, A3 Pro, Icon, Mark, Mark 2, Note, Ray 700, Ray 800, Ray Max, Switch, Turbo, X1, X1 Pro, Z, Z1, Z1 Pro  1. P81 2. Phicomm Passion 660 3. Poco:  * F1, X2  1. Realme:  * 1, 2, 2 Pro, 3, 3 Pro, 3i, 5, 5 Pro, 5i, 5s, C2, C3, U1, X, X2, X2 Pro, XT  1. Samsung Galaxy:  * A10s, A2 Core, A20s, A30s, A5 (2016, 2017), A50s, A7 (2017, 2018), A70, A8, A8 Star, A80, A9 (2018), A9 Pro, C9 Pro, J6, J7 (2017, Duo Max Nxt, Prime Prime 2, Pro 2), J8, M10, M10s, M20, M30s, M30, Note 8, Note 9, Note 10, On Max, On Nxt, On 6, On7 Prime, On 8, On 8 (2018), S20, S20+, S7, S7 Edge, S9, S9+  1. Smartron:  * Srt.phone * t.phone * t.phone P  1. Sony:  * Xperia (R1, R1 Plus, XA1, XA1 Plus, XA1 Ultra, XA2 Ultra, XZ Premium, XZ1, XZ2, Z5 Dual, Z5 Premium Dual)  1. Swipe: 2. Elite Max 3. Elite Plus 4. Tambo TA-2 5. Tecno: 6. Camom (12Air, iClick) 7. i7 8. Phantom 9 9. Spark Power 10. Videocon:   -. ZS5 (Dash, Delite)   1. Vivo: 2. Nex 3. S1, S1 Pro 4. U10, U20 5. V11, V11 Pro, V15, V15 Pro, V17, V17 Pro, V1Max, V3, V3Max, V5, V5 Plus, V5s, V7, V7+, V9, V9 Pro, V9 Youth 6. X21, X7 7. Y12, Y17, Y19,Y66, Y69, Y81, Y82, Y91, Y91i 8. Z1 Pro, Z1x 9. XiaoMi: 10. Redmi:  * 5, 6, 7A, 8, 8A, 8A Dual, Note 5 Pro, Note 7, Note 7 Pro, Note 7S, Note 8, Note 8 Pro, Y1, 3S, 3S Prime, 4, 6 Pro, 7, K20, K20 Pro, Note, Note 3, Note 4, Note 6 Pro, Y2, Y3  1. Mi:  * 4i, A2, A3, Max, Max 2, Mix Max Prime, MIX 2  1. Xolo: 2. 8X-1000 3. Black 4. Black 1X 5. Play 8X-1100 6. Yu: 7. Yunicorn 8. Yureka 9. Yureka (2, Black, Note, Plus) 10. Yutopia 11. Zopo: 12. Flash X Plus 13. Speed 7, 8, X 14. ZTE:  * Blade (A1, A2 Plus, V7 Max, V8 Mini) | 1. Alcatel Pixi 4 (5) 2. Asus Zenfone: 3. 4 4. 6 5. C (ZC451CG) 6. iBall: 7. Andi4.5C Magnifico 8. Andi4 B20 9. Intex: 10. Aqua:  * Air * Air II * Play * Pro4G * Raze * S7 * Strong 5.1 * Wave * Wing Y4  1. Cloud:  * 3G Candy * 3G Gem * Breeze * Champ  1. Karbonn: 2. Alfa A91 power 3. Aura Power 4. Titanium 3-D Plex 5. Lava: 6. A32 7. A52 8. A56 9. Flair P1i 10. Flair S1 11. Iris Fuel F1 Mini 12. LG G Pro Lite 13. Lyf Water 3 14. Micromax: 15. Bharat 3 16. Bolt D303 17. Motorola Moto E 18. Nokia X Dual SIM 19. Onida i4GC1 20. Samsung: 21. Galaxy:  * C7 Pro * Grand Duos * J1 Ace * J2 Ace * Mega 5.8  1. Spice XLife 406 |

Fig 10. Cluster Detail

From the plot and table shown in figure 6 and the table shown in figure 7, we can see that the most significant separator was the number of processors that are present in the phones while the screen sizes are distributed evenly throughout every cluster. As for the table on figure 7, we can see that the majority of phones sold in India were equipped with between 4 to 10 processors which all of them located in cluster 0 and 1. Back to the plot figure 6, we can also see that most of the phones sold in India have 4 and 8 processors. Also using the number of processors as a variable reference for the cluster naming, we can name cluster 0 as “Mid-Range Phone” since the phones inside have 4 to 6 processors and its between cluster 1 and 2; cluster 1 as “High-End Phone” since the phones inside have 6 to 10 processors and its on the top of the plot , and Cluster 2 as “Entry Level Phone” since the phones inside have 1 to 2 processors and its on the bottom of the plot.

1. **Conclusion**

Throughout our study about the utilization of machine learning models, such as PCA (Principal Component Analysis), Outliers filter, and also clustering in classification of different smartphone specifications, we can discover more about the possibilities and potential from the usage of those models to the data processing issue. Along with the results of our practice above from our collected data, we can sum up that the most important part of the phone that influences its performance and types are screen size and the number of processors present inside the phone. We can also conclude that usage of PCA (Principal Component Analysis) is an effective and efficient method of data analysis and visualization where it helps in reducing dimensionality and identifying key patterns.

1. **References**

[1].IBM. (n.d.). Usage of Principal Component Analysis (PCA). Retrieved from <https://www.ibm.com/docs/en/ias?topic=pca-usage>

[2] Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, *36*(2), 451–461. <https://doi.org/10.1016/s0031-3203(02)00060-2>

[3]Ahmar, A. S., Napitupulu, D., Rahim, R., Hidayat, R., Sonatha, Y., & Azmi, M. (2018b). Using K-Means Clustering to Cluster Provinces in Indonesia. *Journal of Physics. Conference Series*, *1028*, 012006. https://doi.org/10.1088/1742-6596/1028/1/012006

(dataset)

*Mobile Phone Specifications and Prices*. (2022, August 14). Kaggle. <https://www.kaggle.com/datasets/pratikgarai/mobile-phone-specifications-and-prices>