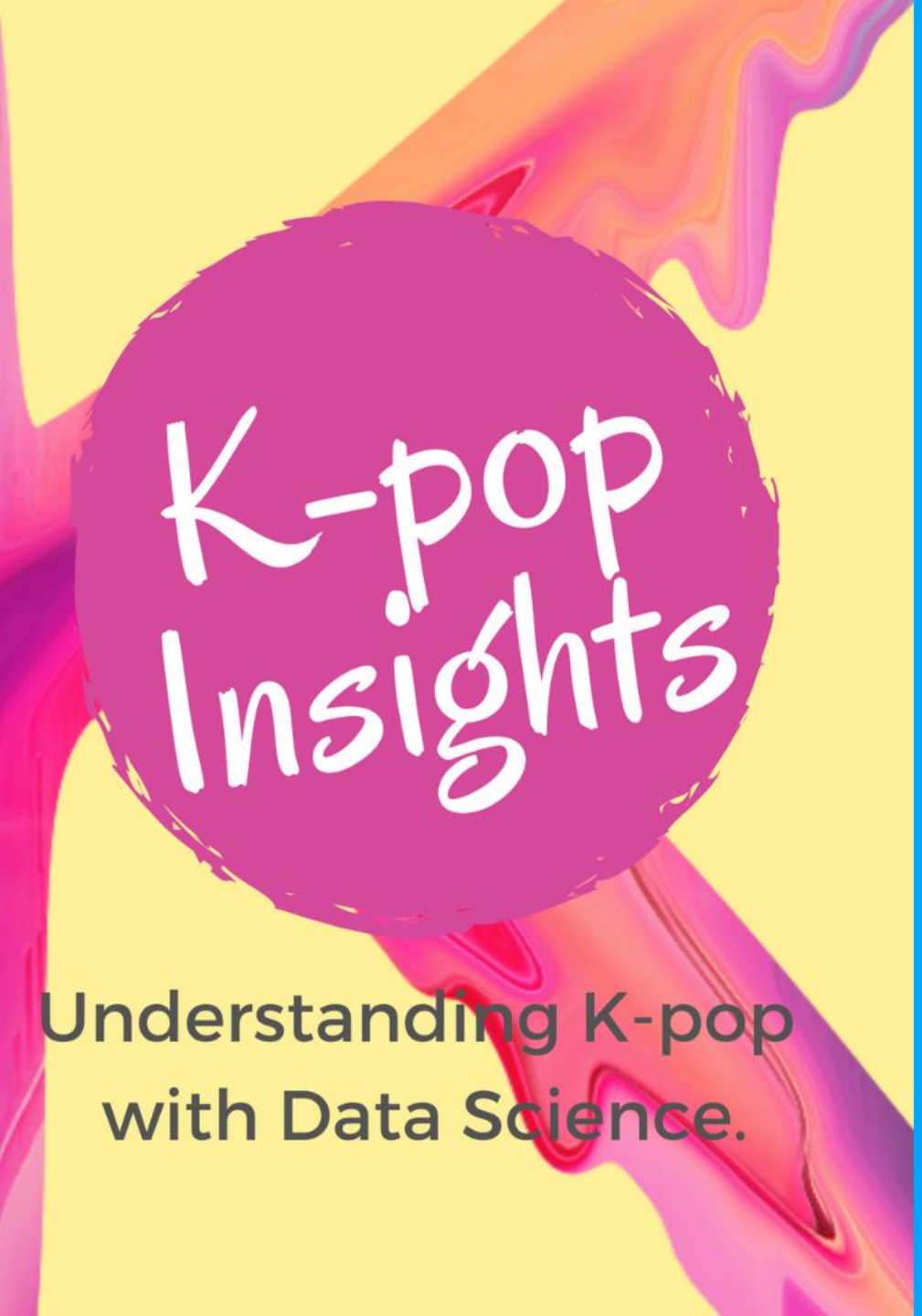


02 - K-pop Groups Segmentation

In this study, I used K-Means Clustering Algorithm to classify K-pop groups into 4 and found some interesting insights using the segmentation .

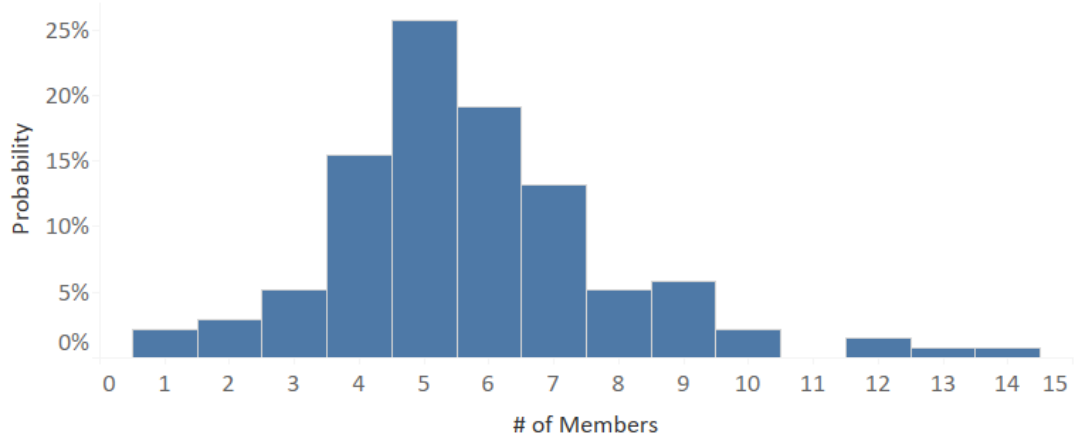
By Yuchen Wang.

Data Source: 2005-2019 kpop data from dbkpop.com

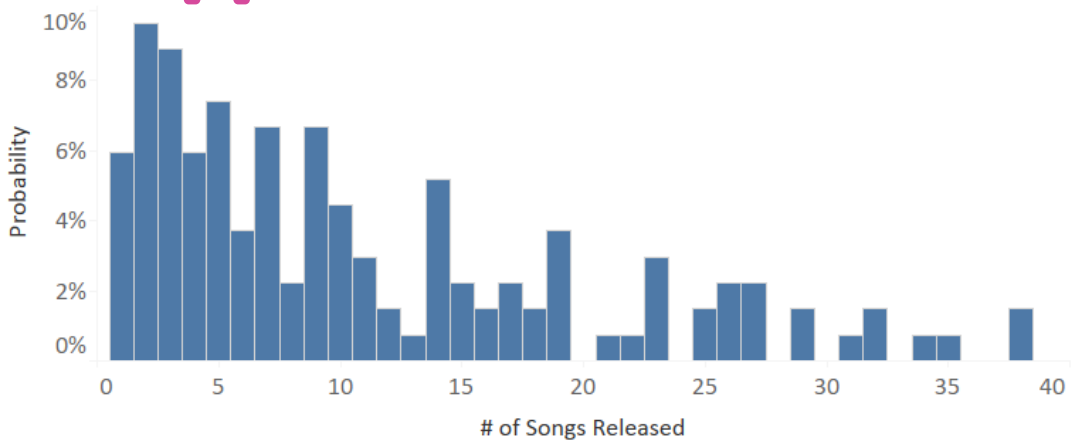
Features Captured

K-pop Insights
Understanding K-pop with Data Science.

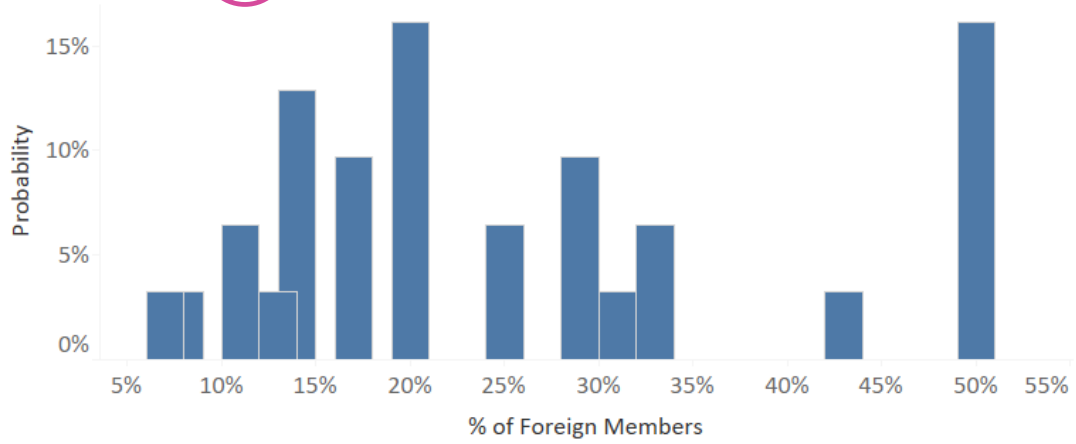
Group Size



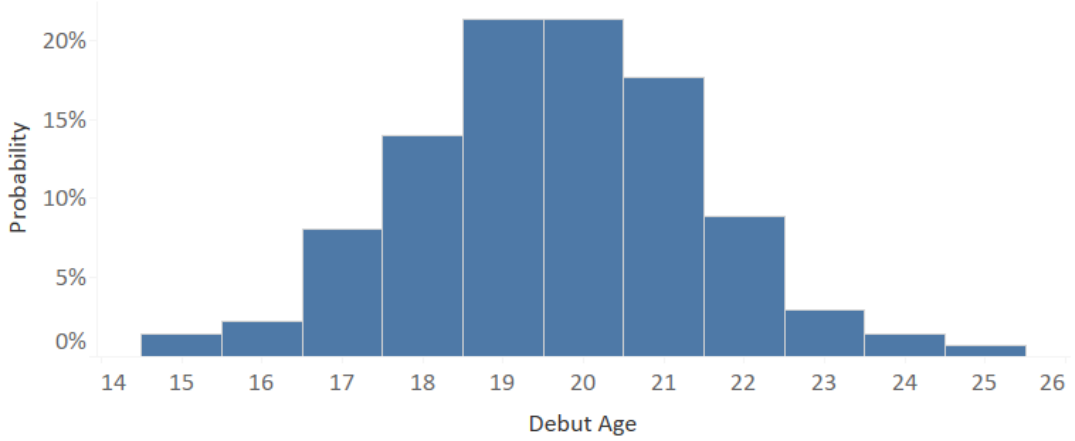
Production



International



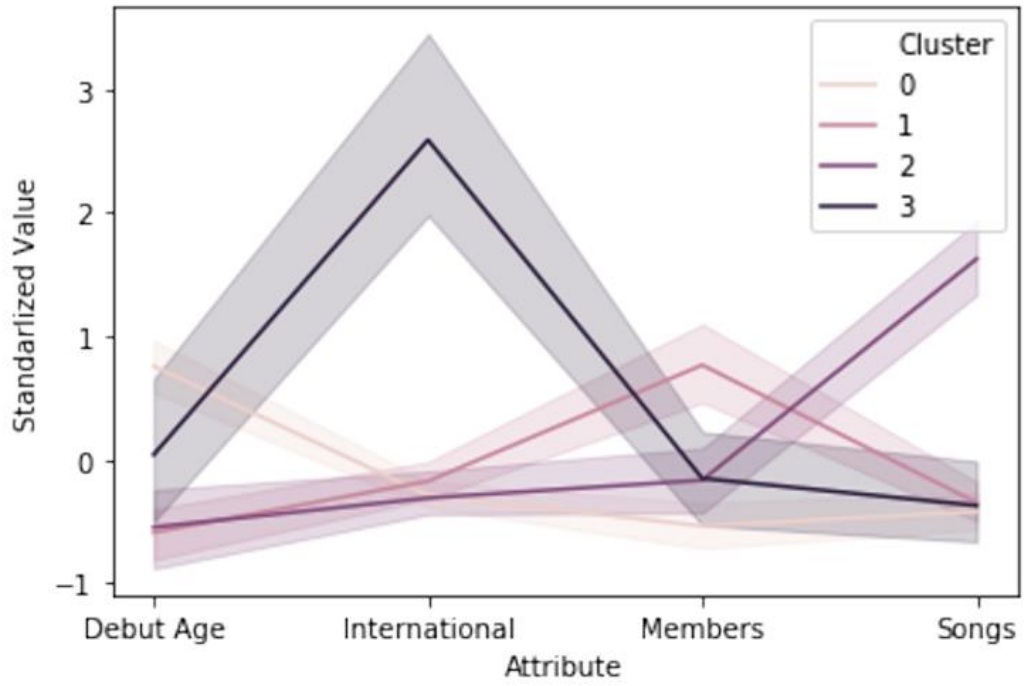
Youthness

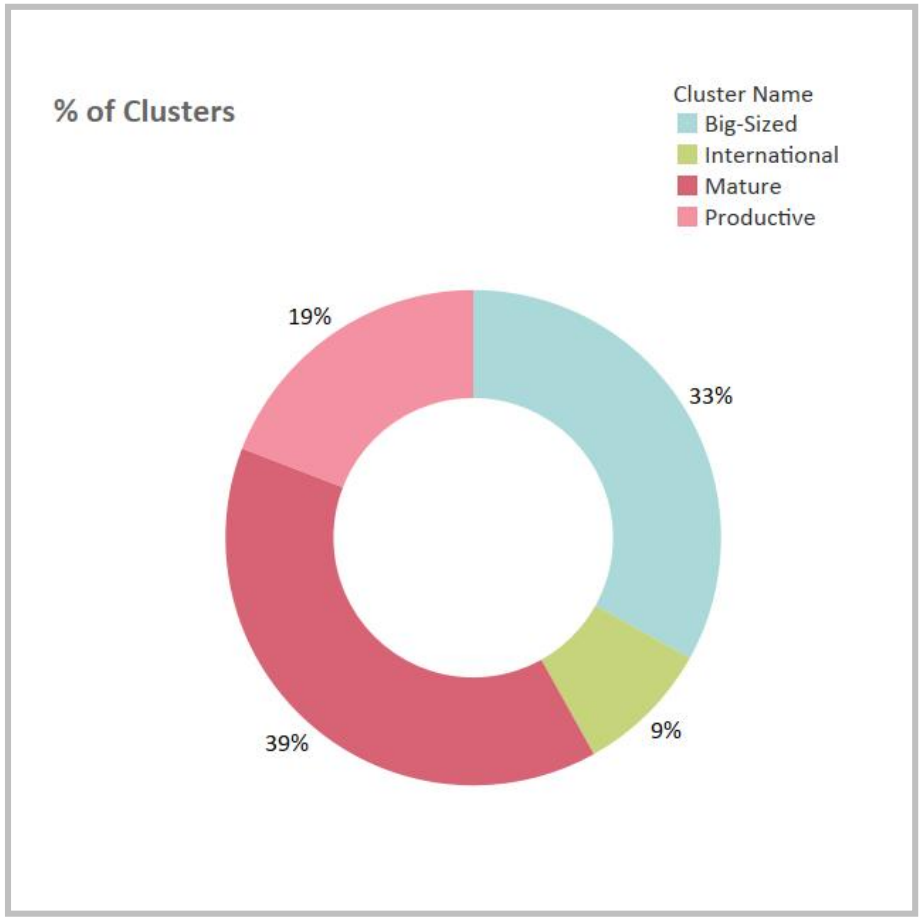


Clustering Output

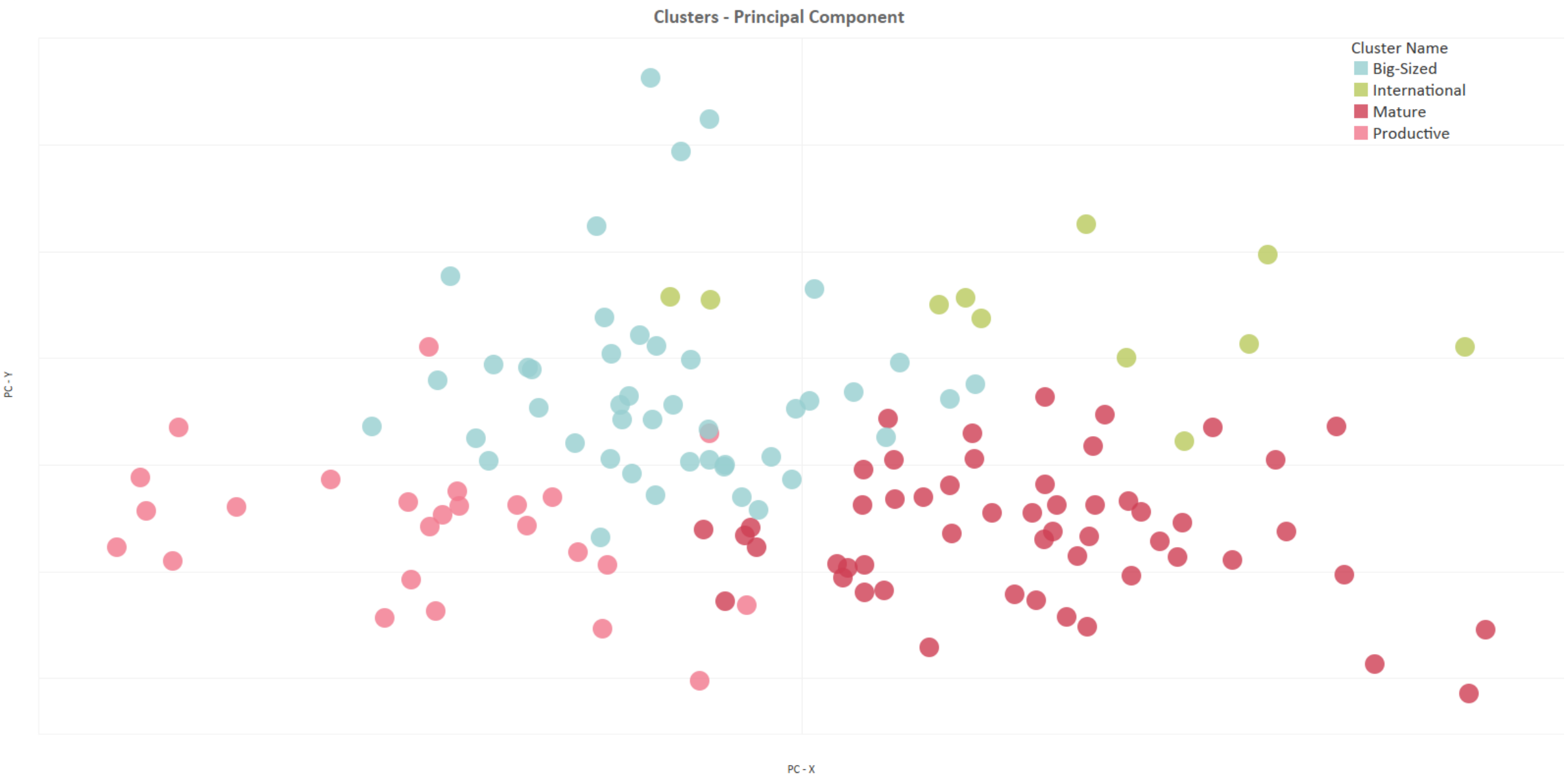
Mean Value

Cluster Name	Avg. Debut Age	% of Foreign Members	# of Members	# of Songs
Big-Sized	19	4%	8	12
International	19	40%	6	14
Mature	21	3%	5	11
Productive	19	2%	6	30





Principal Component Visualization



Key Take-aways

- Features Captured: **4**
- Algorithm: **K-Means Clustering**
- Parameters Tuning: **The Elbow Method**
- 4 Clusters Classified: **Mature, International, Productive, Big-sized**

More in the Future

Some spoiler alerts:

- Dashboards on **PowerBI** and **Google Data Studio**
- Music Videos Sentimental Analysis (**NLP**)
- Music Album Covers Analysis (**Image Processing**)

If you are interested in working on this fun project with me, please feel free to **contact me**!
linkedin.com/in/yuchenwang01

All codes, reports and dashboards are at my **Github**.
github.com/brantgithub/K-pop-Data-Aanalysis

Appendix

- Reference: https://towardsdatascience.com/customer-segmentation-in-python-9c15acf6f945
- Tableau Dashboards: https://public.tableau.com/profile/yuchen.brant.wang
- Reports, Codes & Models: https://github.com/brantgithub/K-pop-Data-Aanalysis