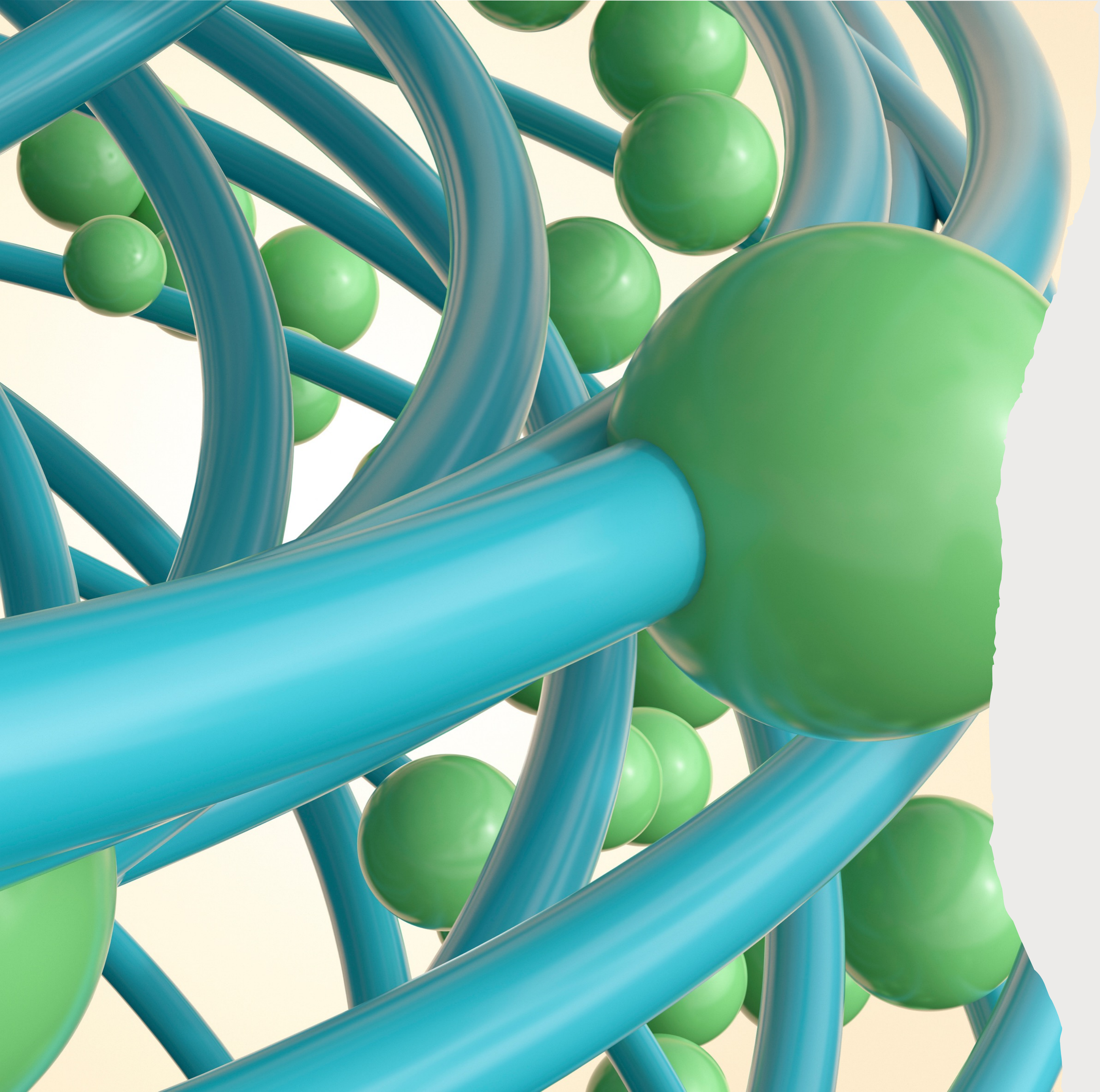
# New Phone. Who’s This?

An Analysis of cell Phone Features and Their Relationship With Pricing

# Why Study Cell Phone Features?

- Used by a Majority of People
- Technological Advances
- Need to Find Best Value for Phone
  - Consumer Purchasing Power
  - Cost to Manufacturers
- Ability to Change Lives
  - Lil Nas X and Tiktok

TikTok



# Related Work

- Nivitus[2]
    - Pandemic Phone Pricing
    - Minimal Attributes
    - No Error Testing
- Al-Shawwa, Abu-Naser and Nasser[3]
    - Neural Network
    - Difficult Interpretability
        - We Want the Opposite
- Asim and Khan[4]
    - Naïve Bayes and Decision Tree Classifiers
    - Inflation Ignored

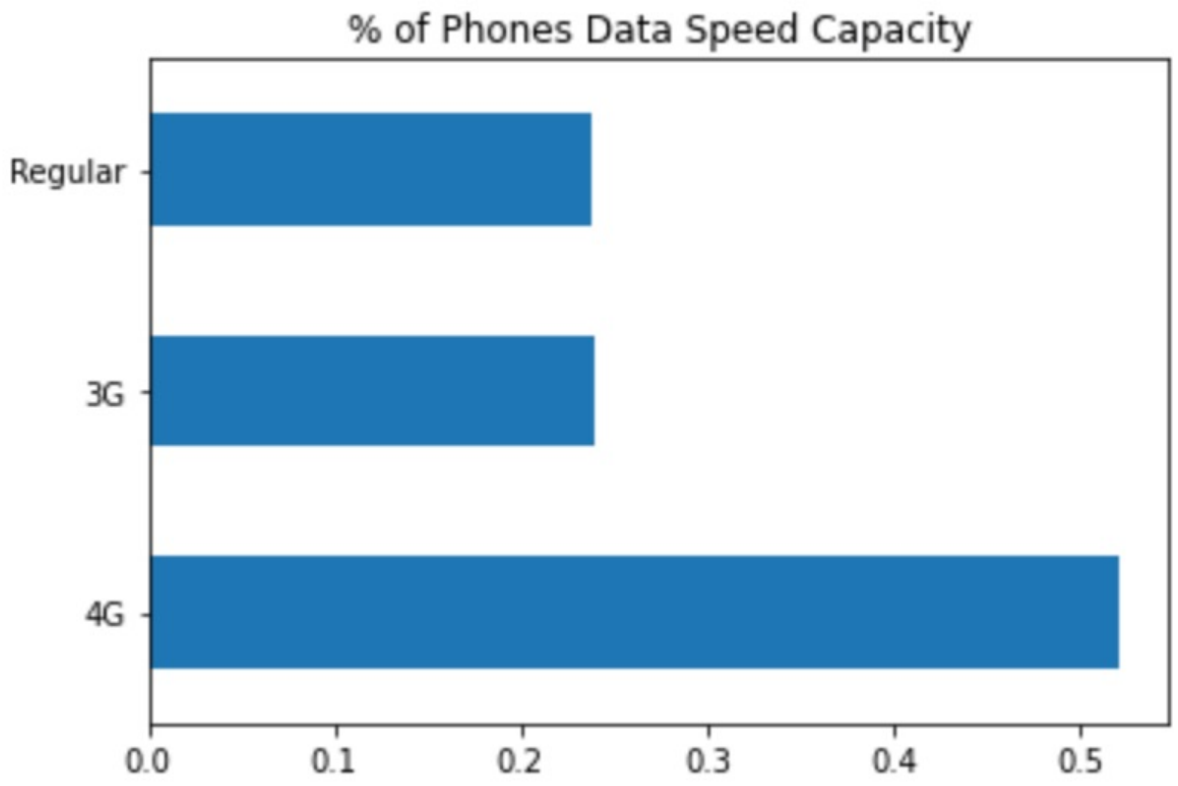
# Proposed Work: Tools and Data Attribute

- Cell Phone Hardware
  - Battery Power
  - Blue
  - Clock Speed
  - Dual Sim
  - Front Camera
  - 4G
  - Internal Memory
  - Mobile Depth
  - Mobile Weight
  - Number of Cores
  - Primary Camera
  - Pixel height
  - Pixel Width
  - RAM
  - Screen Height
  - Screen Width
  - Talk Time
  - 3G
  - Touch Screen
  - Wifi
  - Price Range (Response)

- Analysis Using Python
- Data Set
  - Absihek Sharma
  - Kaggle

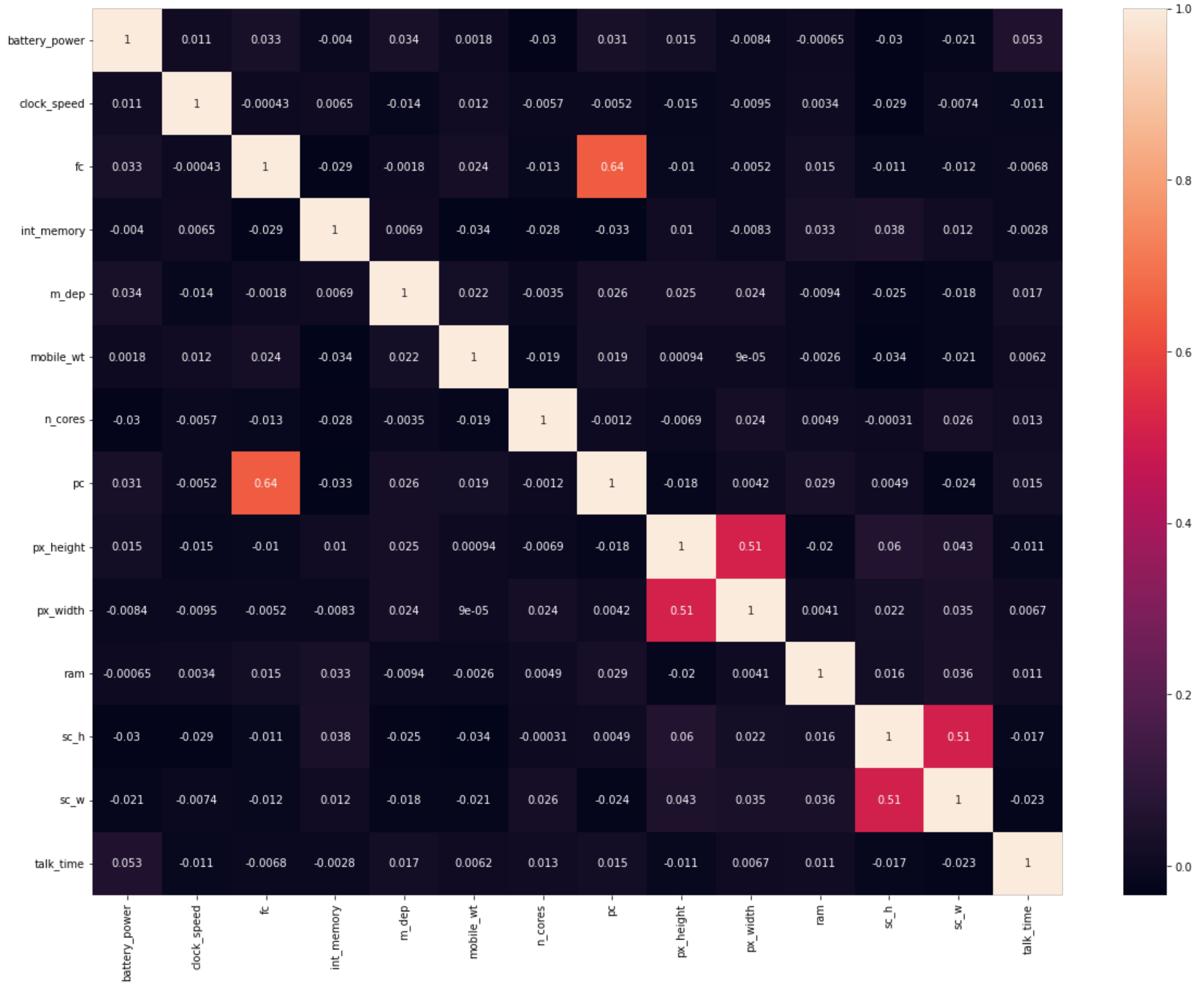
# Data Understanding

- Overview
  - 2000 Observations
  - 14 Continuous and 7 Categorical Variables
  - 7 Binary Categories
- Response Variable
  - Price Range
    - Low
    - Below Average
    - Above Average
    - high
  - 500 Each
- Data Speed Capacity (Graph on Right)
  - Uneven Levels
- Split 70/10/20 for train, validation and tests set
- Metric: Accuracy
  - Want to be as accurate as possible
- Warehousing
  - Two csv files



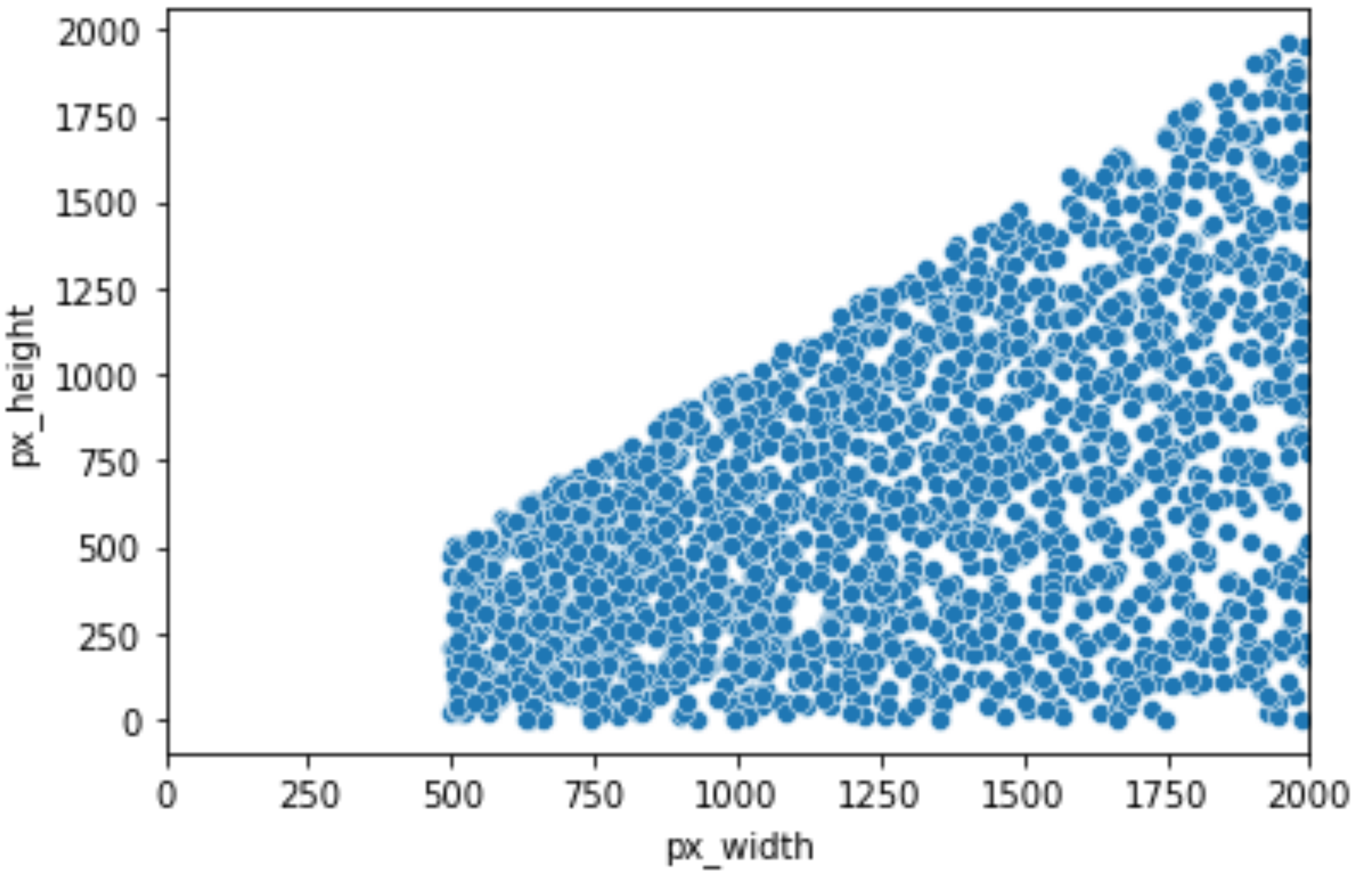
# Continuous Correlation

- Front Camera and Primary Camera Highly Correlated
  - Makes sense for both Cameras to be Similar Quality
- Pixel Width and Pixel Height
  - Expected Correlation



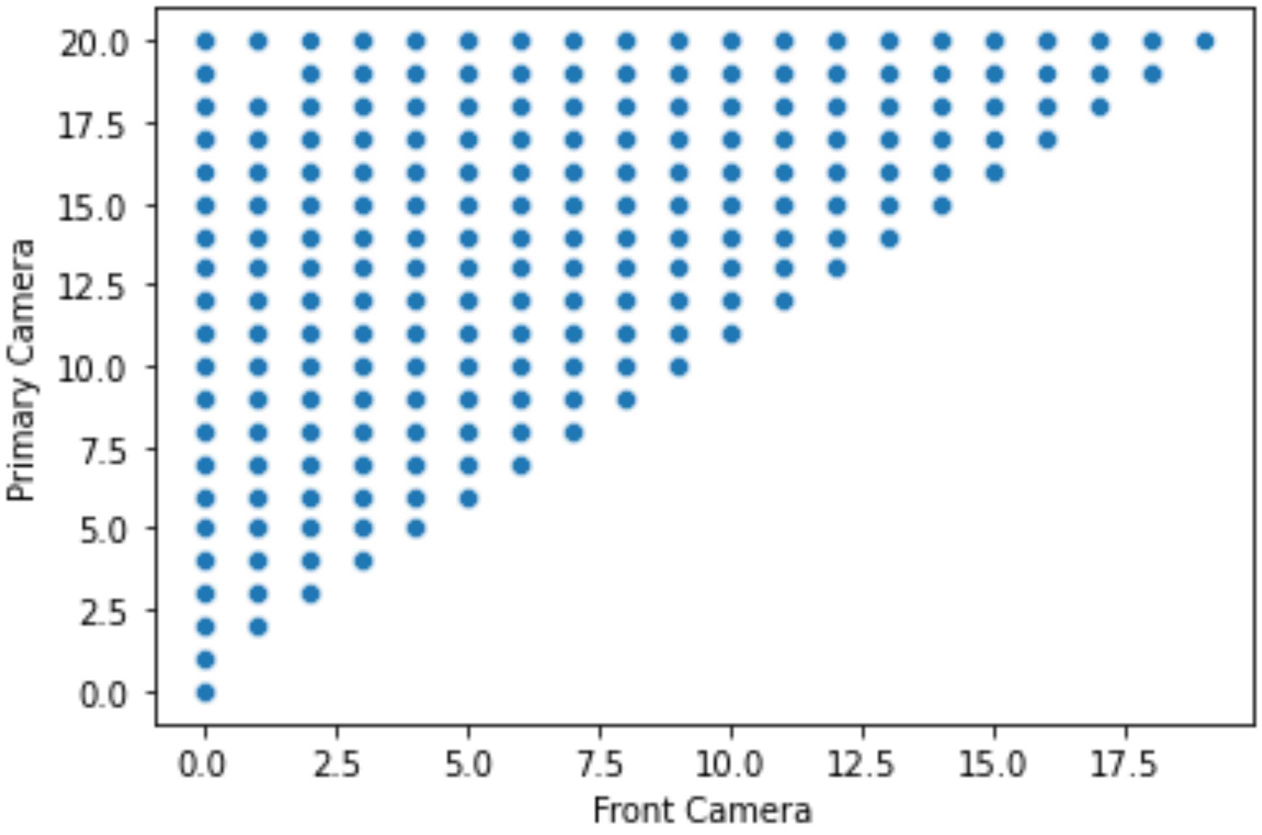
# Pixel Height and Pixel Width

- Minimal Pixel Width is 500
- Height >= Width
- Usual Pixel Resolution
  - 1280 x 1024
  - 1920 x 1080
  - 2560 x 1440



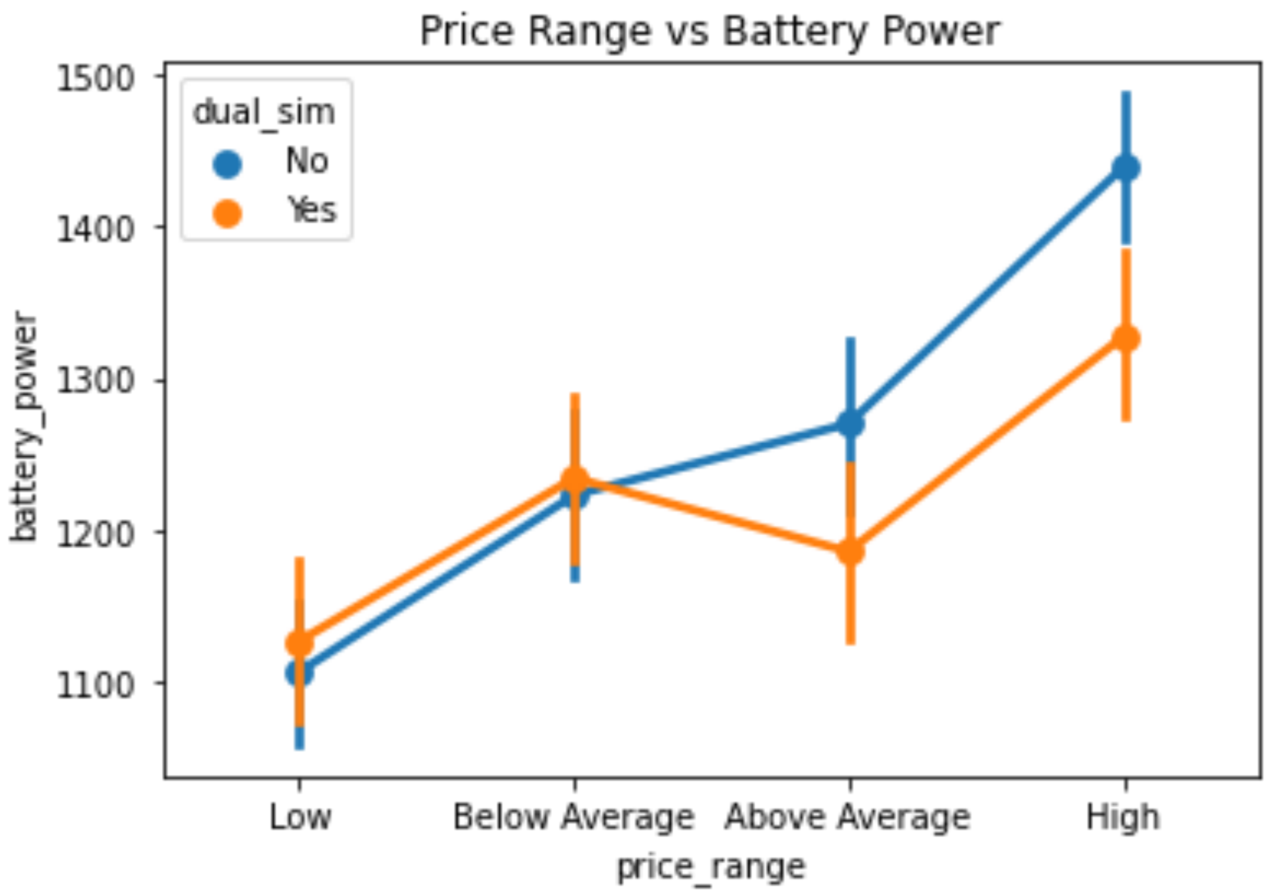
# Price Range vs Battery Power

- Primary Camera >= Front Camera
- Confirms hypothesis Primary Camera > Front Camera



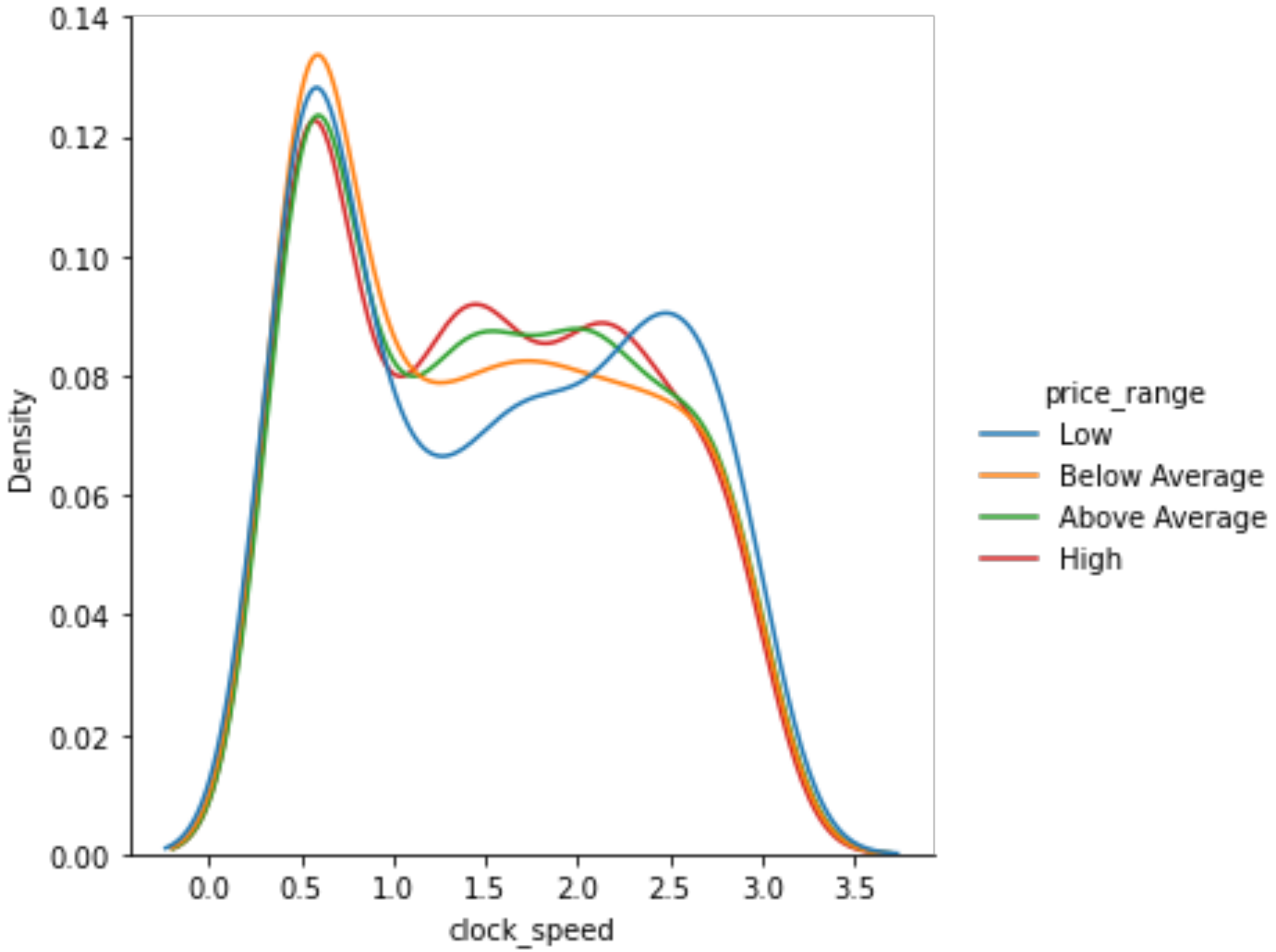
# Primary Camera and Face Camera

- Important, regardless of location
- Dual Sim Cards less  Prevalent in Higher Priced Devices
- Smart Phones Store Data without Sims
- Higher Prices = More Battery Power
- Lower Prices = Less Battery Power



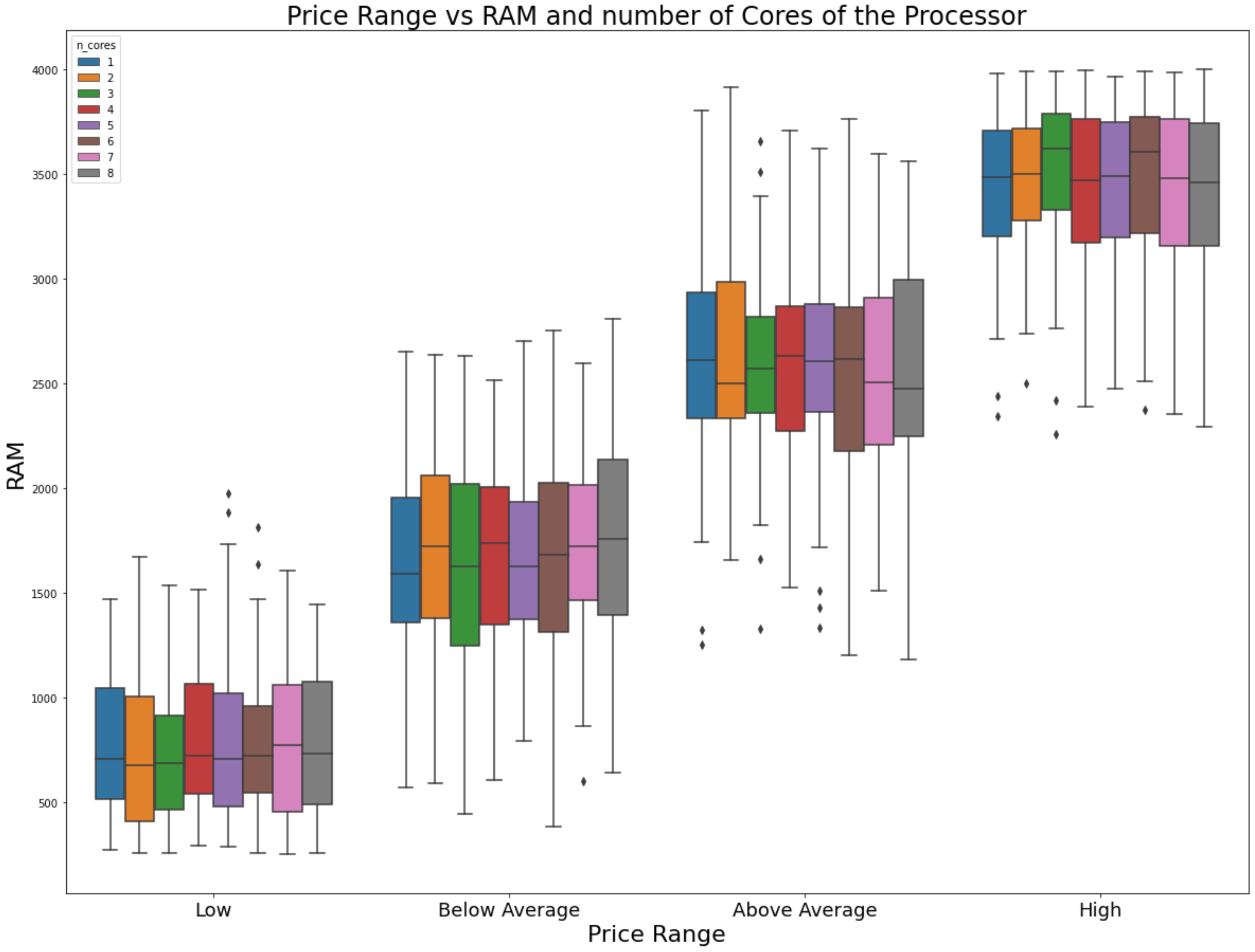
# Price Range and Clock Speed

- Hypothesis: Higher Priced Phones Have Lower Clock Speed
- Low Prices = slower speed
- Other Prices Are Similar Speeds



# Price Range and Clock Speed

- Hypothesis: Higher RAM = More Processors and Expensive Phone
- Visualization Shows No Relationship Between RAM and Processor Count
- The higher the prices, the more RAM



# Price Range and Touch Screen Capability

- Hypothesis: Touch Screen = Expensive Phone
- Hypothesis Test
  - $H_0$: Price Range and Touch Screen Capability are independent
  - $H_1$: Price Range and Touch Screen Capability are not independent.
- P-Value: 0.27
  - Independent

| touch_screen<br>price_range | No | Yes |
|---|---|---|
| Above Average | 265 | 235 |
| Below Average | 239 | 261 |
| High | 252 | 248 |
| Low | 238 | 262 |

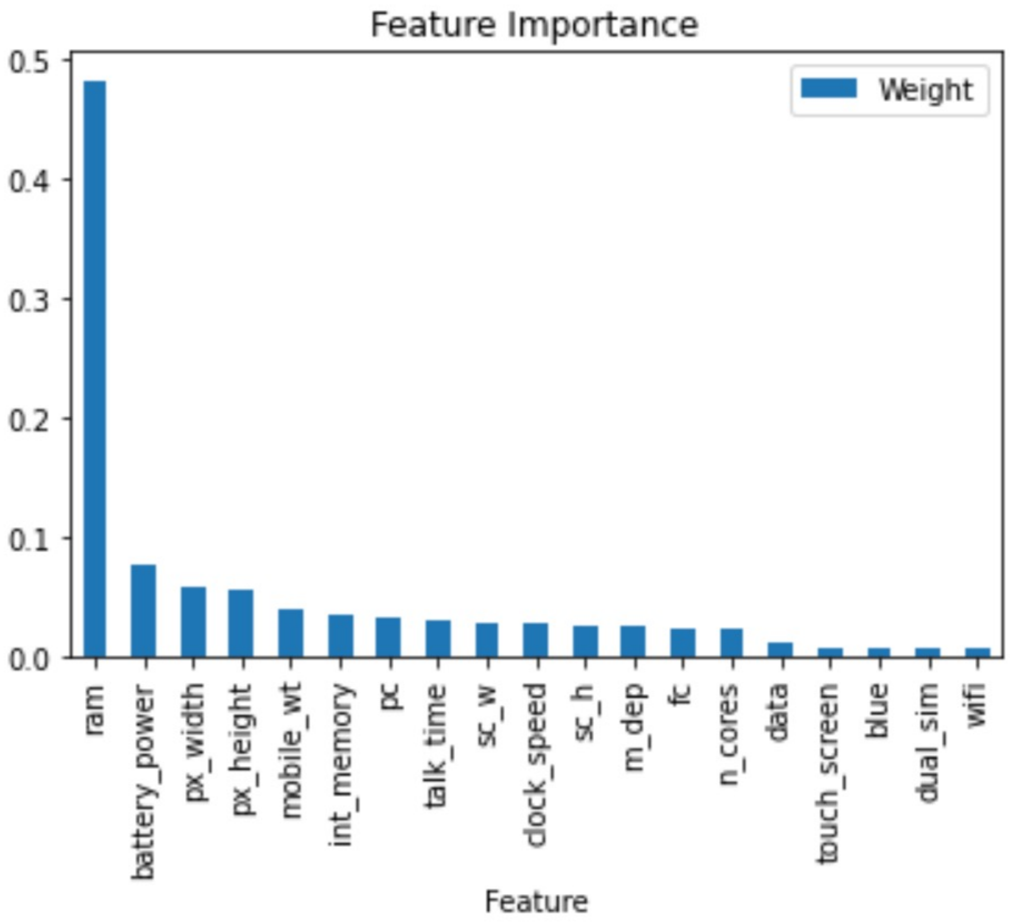
# Model Training and Candidates

- Experimental Method
  - Training Set: 70%
  - Validation Set: 10%
  - Test Set: 20%
- Accuracy: Best Measurement
  - Expensive Phone Deters Consumer and Manufacturer
  - Inexpensive Phone loses profits
- Model Chosen: Random Forest
  - Best Interpretability
  - High Accuracy

| Model | Accuracy (%) |
|---|---|
| Support Vector Machine | 94.56 |
| K Nearest Neighbors | 91.56 |
| Random Forest | 86.69 |
| Decision Tree | 81.69 |
| Multinomial Logistic Regression | 71.00 |

# Feature Importance

- Random Access Memory (RAM)
  - Most Important
- Battery life and Pixel Dimensions contribute a lot
  - Less charging
  - More Mobility
- Dropped Features from Model
  - Data Speed
  - Touch Screen Capability
  - WIFI
  - Bluetooth
  - Dual-Sim



# Validation Set Test

- Accuracy: 89%
- Hyperparameter Tuning: GridSearchCV
  - bootstrap: True
  - criterion: Entropy
  - max_Depth: 10
  - min_samples_split: 5
  - n_estimators: 500
- Precision and Recall Improvement

**Before Hyperparameter Tuning**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Low | 0.88 | 0.93 | 0.91 | 88 |
| Below Avg | 0.88 | 0.86 | 0.87 | 108 |
| Above Avg | 0.83 | 0.88 | 0.85 | 97 |
| High | 0.95 | 0.88 | 0.91 | 107 |
| accuracy | | | 0.89 | 400 |
| macro avg | 0.89 | 0.89 | 0.89 | 400 |
| weighted avg | 0.89 | 0.89 | 0.89 | 400 |

**After Hyper Parameter Tuning**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Low | 0.88 | 0.94 | 0.91 | 88 |
| Below Avg | 0.88 | 0.86 | 0.87 | 108 |
| Above Avg | 0.86 | 0.86 | 0.86 | 97 |
| High | 0.94 | 0.92 | 0.93 | 107 |
| accuracy | | | 0.89 | 400 |
| macro avg | 0.89 | 0.89 | 0.89 | 400 |
| weighted avg | 0.89 | 0.89 | 0.89 | 400 |

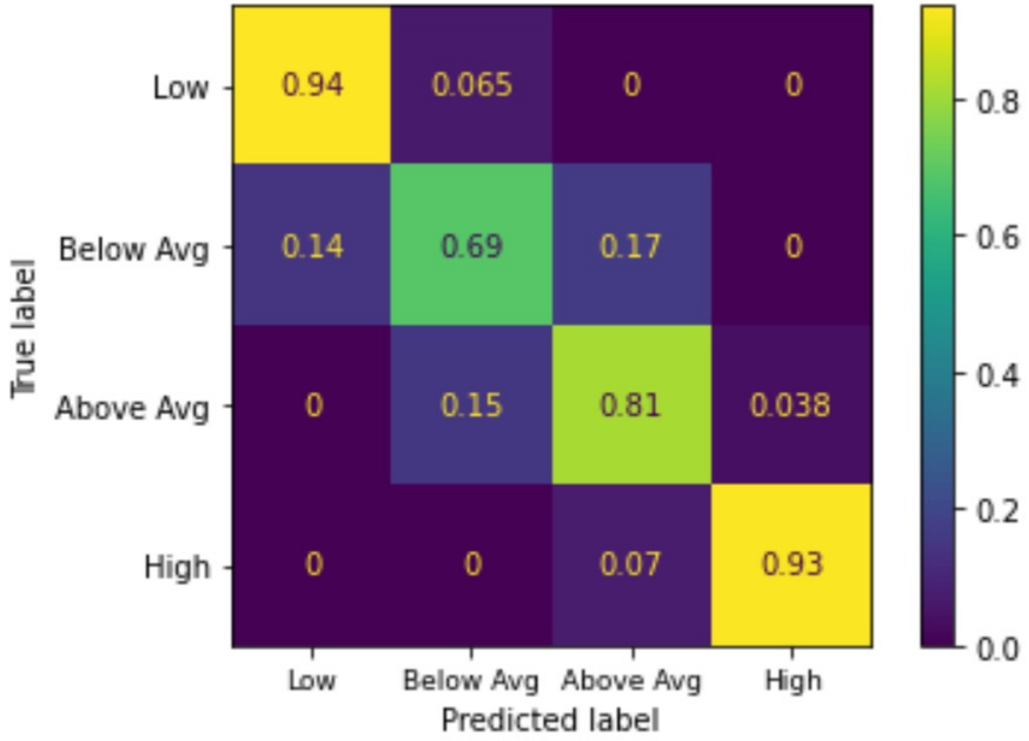
# Final Model

- Accuracy Decreased a small amount
  - Minimal Overfitting
- Low Precision/Recall: Below-Avg
  - F1 score
  - Class imbalance issue?
- Indistinguishable below-avg and above avg?

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Low | 0.91 | 0.94 | 0.92 | 62 |
| Below Avg | 0.71 | 0.69 | 0.70 | 42 |
| Above Avg | 0.81 | 0.81 | 0.81 | 53 |
| High | 0.95 | 0.93 | 0.94 | 43 |
| accuracy | | | 0.85 | 200 |
| macro avg | 0.84 | 0.84 | 0.84 | 200 |
| weighted avg | 0.85 | 0.85 | 0.85 | 200 |

# Evaluation

- Below average not as accurate
- Below Avg and Above Avg mistaken for each other
- Too Many Similarities?



# Discussion

- Timeline Completed
  - January 17: Have Proposal Completed
  - January 24: Have all data collected
  - February 2: Finish Cleaning and Combining Datasets
  - February 15: Finish all Preprocessing and Outstanding Visualizations
  - February 19: Finish Comparing Potential Models
  - February 22: Tune Final Model and Test
  - February 27: Submit project with presentation
- Challenge: Time
  - Completed within 8 week
  - Would like more features
- Future work
  - Brand
  - Temporal Changes
  - Software

# Conclusion

- Cell Phone Evolution
- Customers and Manufacturers Need Smart Decisions
- Best Value

# References


- [1] G., Deyan. (2022, February 6). *How Much Time Does The Average American Spend on Their Phone in 2022?*. Techjury. https://techjury.net/blog/how-much-time-does-the-average-american-spend-on-their-phone/#gref
- [2] Nivitus. (2020, July 22). *Mobile Price Prediction Using Machine Learning.* Medium. https://medium.com/@Nivitus./mobile-price-prediction-using-machine-learning-fa9cab6fb242
- [3] Al-Shawwa, Mohammed; Abu-Naser, Samy S. and Mohammed Nasser, Ibrahim. (February 2019). *Developing Artificial Neural Network for Predicting Mobile Phone Price Range.* ResearchGate. https://www.researchgate.net/publication/331398317_Developing_Artificial_Neural_Network_for_Predicting_Mobile_Phone_Price_Range
- [4] Asim, Muhammad and Khan, Zafar. (2018, March). *Mobile Price Class Prediction Using Machine Learning Techniques.* Towards Data Science. https://www.ypulse.com/article/2019/10/01/how-gen-z-millennials-are-discovering-music-on-social-media-in-2-charts/
- [5] Zeman, Eric. (2020, July). *Phone Prices Have Almost Doubled Over the Past Five Years. Why is That?.* Android Authority. https://www.androidauthority.com/how-smartphone-prices-have-changed-1134574/
- [6] Sharma, Abhishek. *Mobile Price Classification*. Kaggle. https://www.kaggle.com/iabhishekofficial/mobile-price-classification