Summary Presentation

The App Project

- Project Overview:

XYZ Company designs operating systems, and they want to build a major apps store into their user interface. To this end, they want to know whether Google Play apps have higher reviews on average than Apple Store apps (or vice versa), as they're intending to strike a deal with just one of these companies.

Brief Explanation of Major Steps
1) Sourcing and loading data

Google

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19M	10,000+	Free	0	Everyone	Art & Design	January 7, 2018	1.0.0	4.0.3 and up
1	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500,000+	Free	0	Everyone	Art & Design;Pretend Play	January 15, 2018	2.0.0	4.0.3 and up
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Design	August 1, 2018	1.2.4	4.0.3 and up

Apple

	Unnamed: 0	id	track_name	size_bytes	currency	price	rating_count_tot	rating_count_ver	user_rating	user_rating_ver	ver	cont_rating	prime_genr
0	1	281656475	PAC-MAN Premium	100788224	USD	3.99	21292	26	4.0	4.5	6.3.5	4+	Game
1	2	281796108	Evernote - stay organized	158578688	USD	0.00	161065	26	4.0	3.5	8.2.2	4+	Productivi
2	3	281940292	WeatherBug - Local Weather, Radar, Maps, Alerts	100524032	USD	0.00	188583	2822	3.5	4.5	5.0.0	4+	Weath

Brief Explanation of Major Steps
2) Cleaning and transforming data

Google

```
prime_genre            object
user_rating           float64
rating_count_tot        int64
price                 float64
dtype: object    array([  3.99,   0.  ,   0.99,   9.99,   4.99,   7.99,   2.99,   1.99,
                          5.99,  12.99,  21.99, 249.99,   6.99,  74.99,  19.99,   8.99,
                         24.99,  13.99,  14.99,  16.99,  47.99,  11.99,  59.99,  15.99,
                         27.99,  17.99, 299.99,  49.99,  23.99,  20.99,  39.99,  99.99,
                         29.99,  34.99,  18.99,  22.99])
```
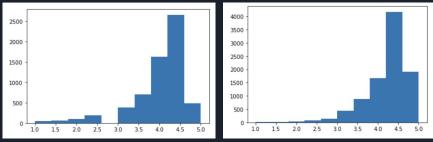
Apple

```
Category     object
Rating      float64
Reviews      object
Price        object
dtype: object    array(['0', '$4.99', '$3.99', '$6.99', '$1.49', '$2.99', '$7.99', '$5.99',
                        '$3.49', '$1.99', '$9.99', '$7.49', '$0.99', '$9.00', '$5.49',
                        '$10.00', '$24.99', '$11.99', '$79.99', '$16.99', '$14.99',
                        '$1.00', '$29.99', '$12.99', '$2.49', '$10.99', '$1.50', '$19.99',
                        '$15.99', '$33.99', '$74.99', '$39.99', '$3.95', '$4.49', '$1.70',
                        '$8.99', '$2.00', '$3.88', '$25.99', '$399.99', '$17.99',
                        '$400.00', '$3.02', '$1.76', '$4.84', '$4.77', '$1.61', '$2.50',
                        '$1.59', '$6.49', '$1.29', '$5.00', '$13.99', '$299.99', '$379.99',
                        '$37.99', '$18.99', '$389.99', '$19.90', '$8.49', '$1.75',
                        '$14.00', '$4.85', '$46.99', '$109.99', '$154.99', '$3.08',
                        '$2.59', '$4.80', '$1.96', '$19.40', '$3.90', '$4.59', '$15.46',
                        '$3.04', '$4.29', '$2.60', '$3.28', '$4.60', '$28.99', '$2.95',
                        '$2.90', '$1.97', '$200.00', '$89.99', '$2.56', '$30.99', '$3.61',
                        '$394.99', '$1.26', 'Everyone', '$1.20', '$1.04'], dtype=object)
```

- Cleaned and transformed Google/Apple data
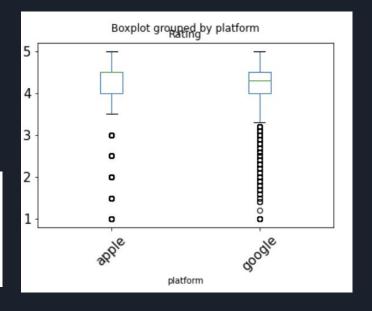- Join two datasets

	Category	Rating	Reviews	Price	platform
0	ART_AND_DESIGN	4.1	159	0.00	google
1	ART_AND_DESIGN	3.9	967	0.00	google
2	ART_AND_DESIGN	4.7	87510	0.00	google
3	ART_AND_DESIGN	4.5	215644	0.00	google
4	ART_AND_DESIGN	4.3	967	0.00	google
...	...	...	...	...	...
18032	Games	4.5	142	0.00	apple
18033	Games	4.5	30	0.00	apple
18034	Utilities	4.5	15	1.99	apple
18035	Games	4.5	85	0.00	apple
18036	Games	5.0	3	0.00	apple

15634 rows × 5 columns

Brief Explanation of Major Steps
3) Visualizing and modeling data

Apple

Google

platform	count	mean	std	min	25%	50%	75%	max
apple	6268.0	4.049697	0.726943	1.0	4.0	4.5	4.5	5.0
google	9366.0	4.191757	0.515219	1.0	4.0	4.3	4.5	5.0

- **Brief Explanation of Major Steps**
  3) Visualizing and modeling data

Permutation test
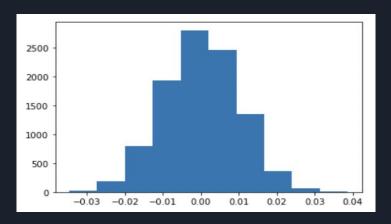
platform	count	mean	std	min	25%	50%	75%	max
apple	6268.0	4.127696	0.625440	1.0	4.0	4.3	4.5	5.0
google	9366.0	4.139558	0.604439	1.0	4.0	4.3	4.5	5.0

Previous analytical summary

platform	count	mean	std	min	25%	50%	75%	max
apple	6268.0	4.049697	0.726943	1.0	4.0	4.5	4.5	5.0
google	9366.0	4.191757	0.515219	1.0	4.0	4.3	4.5	5.0

Difference: **0.14206054745123442**

- Findings

➢ Zero differences are at least as extreme as our observed difference!

➢ The p-value of our observed data is 0.

➢ It doesn't matter which significance level we pick; our observed data is statistically significant, and we reject the Null.

➢ We conclude that platform does impact on ratings. Specifically, we should advise our client to integrate **only Google Play** into their operating system interface.