YOUTUBE ANALYSIS

STATE OF MEDIA 2019 DIGITAL VIDEO

235M

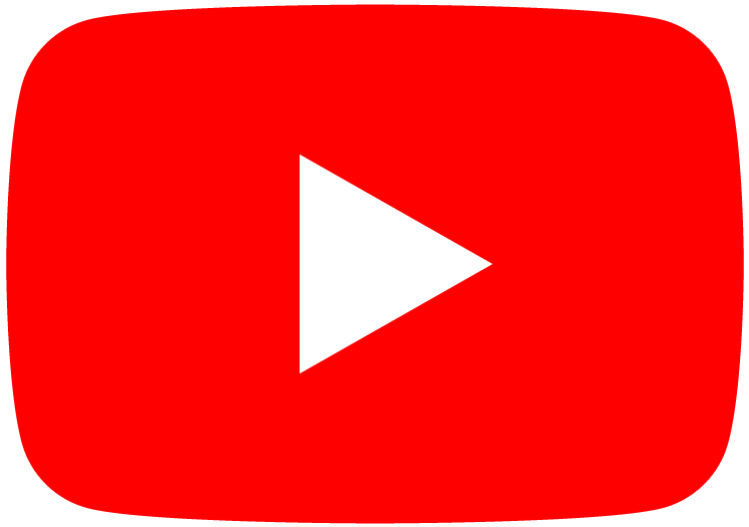
U.S. ACTIVE DIGITAL VIDEO VIEWERS

71% OF THE U.S. POPULATION

+18M NEXT 4 YEARS

Source: eMarketer March 2019

YOUTUBE: THE LEADER IN WATCH TIME



1.90B

GLOBAL USERS

+1.0B

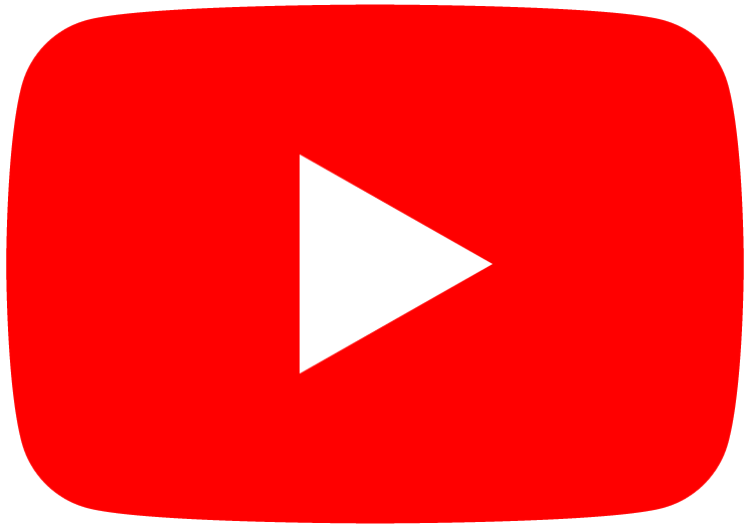
DAILY HOURS WATCHED

70%

WATCH TIME FROM MOBILE

Source: YouTube For Press, 2019

YOUTUBE WILL ADD 13M NEW U.S. USER BY 2023



197M

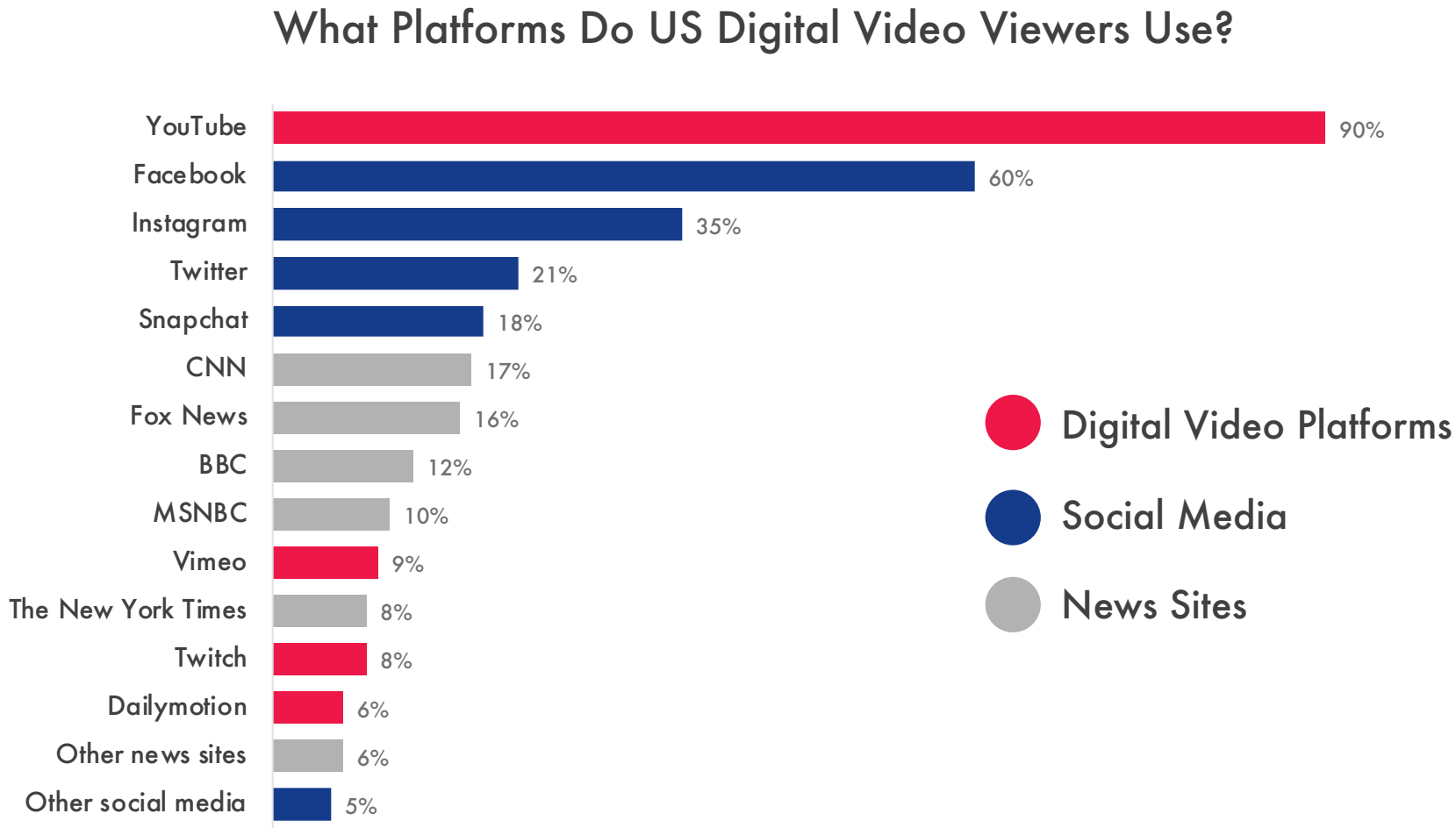
U.S. MONTHLY ACTIVE VIEWERS

83%

OF TOTAL U.S DIGITAL VIDEO VIEWERS

Source: eMarketer March 2019

SOCIAL MEDIA PLATFORMS ARE BIG VIDEO DRIVERS



Source: AudienceProject Q4, Jan 2019

YOUTUBE DEEP DIVE

EXECUTING OUR ANALYTICS MODEL

IT STARTS WITH A BUSINESS QUESTION

YOUTUBE CHANNEL GROWTH

Objective: Discover What Impacts YouTube
Subscriber Growth

PROCESS

WHAT WAS THE PROCESS FOR ANSWERING THE GROWTH QUESTION ?

We used **statistics** and **predictive modeling** to identify what impacts subscriber growth

REVIEW FINDINGS

YOUTUBE CHANNEL GROWTH

Statistics identified the following:

That there are **6 metrics** directly related to subscriber growth

- Shares
- V - 7 Views (Paid & Organic)
- Total Completed Views
- V - 7 Minutes Watched (Paid & Organic)
- Total Ad Views V-28
- Comments

JUPYTER NOTEBOOK TRY 1

In [1]:
```
import numpy as np
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
```

In [2]:
```
data = pd.read_csv('/Users/ncabayan/Desktop/YTGlobal.csv')
```

In [3]:
```
y = data['Subs Gained']
x1 = data[['Video Length', 'Likes', 'Dislikes', 'Comments','Shares', 'Total Engagements', 'Total Views',
           'Total Ad Views Paid','Total Organic Views', 'Total Views 7 days after published', 'Total Ad Views View 7',
           'Total Organic Views View 7','Total Views 28 days after published','Total Ad Views View 28',
           'Total Organic Views View 28','Total Completed Views','Total Minutes Watched',
           'Total Ad Minutes Watched','Total Organic Minutes Watched',
           'Total Minutes Watched 7 days after published','Total Ad Minutes Watched 7 days after published',
           'Total Organic Minutes Watched 7 days after published','Total Minutes Watched 28 days after published',
           'Total Ad Minutes Watched 28 days after published',
           'Total Organic Minutes Watched 28 days after published', 'Average View Duration',
           'Avg View Duration In Minutes','Avg Percent Viewed']]
```

In [4]:
```
x = sm.add_constant(x1)
results = sm.OLS(y,x).fit()
results.summary()
```

Out[4]:

OLS Regression Results

Dep. Variable:	Subs Gained	**R-squared:**	0.959
Model:	OLS	**Adj. R-squared:**	0.958
Method:	Least Squares	**F-statistic:**	656.9
Date:	Fri, 08 Feb 2019	**Prob (F-statistic):**	0.00
Time:	08:43:25	**Log-Likelihood:**	-4590.1
No. Observations:	582	**AIC:**	9222.
Df Residuals:	561	**BIC:**	9314.
Df Model:	20		
Covariance Type:	nonrobust		

JUPYTER NOTEBOOK TRY 2

In [1]:
```
import numpy as np
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
```

In [2]:
```
data = pd.read_csv('/Users/ncabayan/Desktop/YTGlobal.csv')
data
```

Out[2]:

	Video Length	Likes	Dislikes	Comments	Shares	Total Engagements	Total Views	Total Ad Views Paid	Total Organic Views	Total Views 7 days after published	...	Total Ad Minutes Watched 7 days after published	Total Organic Minutes Watched 7 days after published	Total Minutes Watched 28 days after published	Total Ad Minutes Watched 28 days after published	Tota Organic Minutes Watched 28 days after published
0	552	1345	29	80	293	1747	59634	4	59630	47945	...	16	141817	193784	16	193768
1	151	522	11	30	76	639	18669	0	18669	14974	...	0	18196	23884	0	23884
2	482	484	20	48	55	607	21114	6	21108	19074	...	3	37856	45495	3	45492
3	173	854	12	47	216	1129	27786	0	27786	20009	...	0	37471	54244	0	54244
4	144	786	15	54	104	959	19908	0	19908	17492	...	0	23421	27006	0	27006
5	140	687	28	34	62	811	27130	0	27130	23111	...	0	26612	31824	0	31824
6	545	1353	56	97	224	1730	78104	11	78093	65375	...	0	175202	232672	4	232668
7	190	887	21	57	140	1105	43763	0	43763	35742	...	0	57063	70593	0	70593

In [3]:
```
# Varriable list
```

In [4]:
```
y = data['Subs Gained']
x1 =data['Shares']
x2 =data['Total Views 7 days after published']
x3 =data['Total Minutes Watched 7 days after published']
x4 =data['Total Ad Views View 28']
x5 =data['Total Completed Views']
x6 =data['Comments']
```

In [5]:
```
# X1 is Shares
```

JUPYTER NOTEBOOK SHARES

In [1]:
```
import numpy as np
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
```

In [2]:
```
data = pd.read_csv('/Users/ncabayan/Desktop/Shares.csv')
```

In [3]:
```
y = data['Subs Gained']
x1 = data[['Video Length', 'Likes', 'Dislikes', 'Comments','Shares', 'Total Engagements', 'Total Views',
           'Total Ad Views Paid','Total Organic Views', 'Total Views 7 days after published', 'Total Ad Views View 7',
           'Total Organic Views View 7','Total Views 28 days after published','Total Ad Views View 28',
           'Total Organic Views View 28','Total Completed Views','Total Minutes Watched',
           'Total Ad Minutes Watched','Total Organic Minutes Watched',
           'Total Minutes Watched 7 days after published','Total Ad Minutes Watched 7 days after published',
           'Total Organic Minutes Watched 7 days after published','Total Minutes Watched 28 days after published',
           'Total Ad Minutes Watched 28 days after published',
           'Total Organic Minutes Watched 28 days after published', 'Average View Duration',
           'Avg View Duration In Minutes','Avg Percent Viewed']]
```

In [4]:
```
x = sm.add_constant(x1)
results = sm.OLS(y,x).fit()
results.summary()
```

Out[4]:

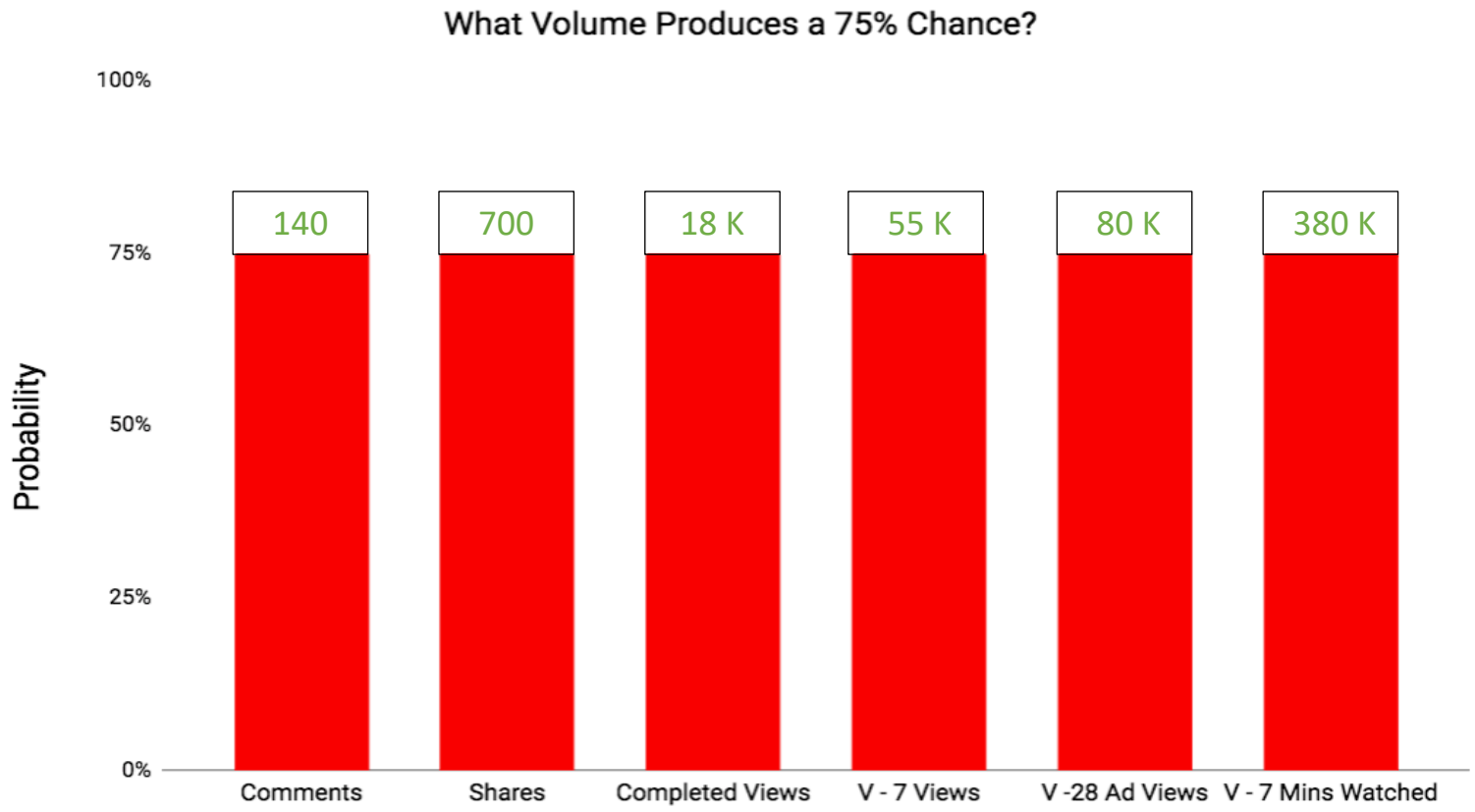
OLS Regression Results

Dep. Variable:	Subs Gained	**R-squared:**	0.838
Model:	OLS	**Adj. R-squared:**	0.832
Method:	Least Squares	**F-statistic:**	128.5
Date:	Tue, 12 Feb 2019	**Prob (F-statistic):**	1.57e-181
Time:	14:09:41	**Log-Likelihood:**	-3544.0
No. Observations:	517	**AIC:**	7130.
Df Residuals:	496	**BIC:**	7219.
Df Model:	20		
Covariance Type:	nonrobust		

REVIEW FINDINGS

PREDICTIVE MODELING IDENTIFIED THE FOLLOWING:

That a post need to hit at least **one** of the volumes below



YOUTUBE DEEP DIVE INSIGHTS

THINGS TO PRIORITIZE:

- The first 7 days after an asset is published
- Boosting between 7 to 28 days
- Get users to complete videos
- Finding ways to gain shares
- Thinking of ways to encourage comments


THANK YOU

QUESTIONS

?