



Facebook Data Analysis

AVANI SHARMA

3 SEPTEMBER, 2018

DATA SOURCE: <https://www.kaggle.com/sheenabatra/facebook-data>

Variables

Age
dob_day
dob_year
dob_month
gender
tenure
friend_count
friendships_initiated
likes / likes received
mobile likes / likes received
www likes / likes received



Exploratory Data Analysis : Key Insights

Gender Interaction Women interact more with fb then men

- Women receive & give more likes than men on average
- Women initiate less friendships than men compared proportionally to friend count

Likes Split Up Shows inclination towards mobile apps

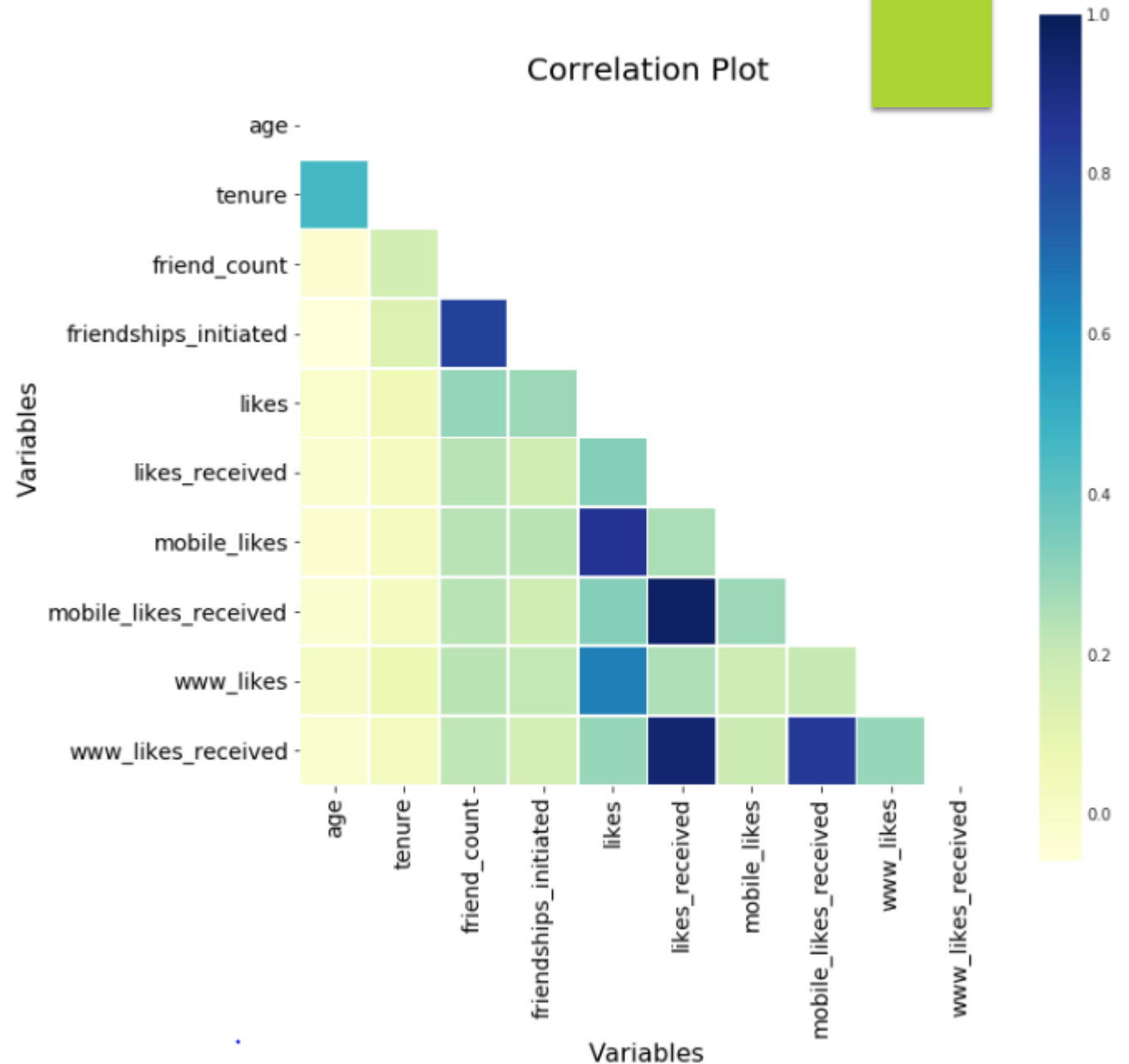
- More prominence of likes from mobile compared to site though few users still interact with sites
- There can be a gradual shift from mobile to site in years to come seeing the trend

User Counts Bump between 1940 to 1980

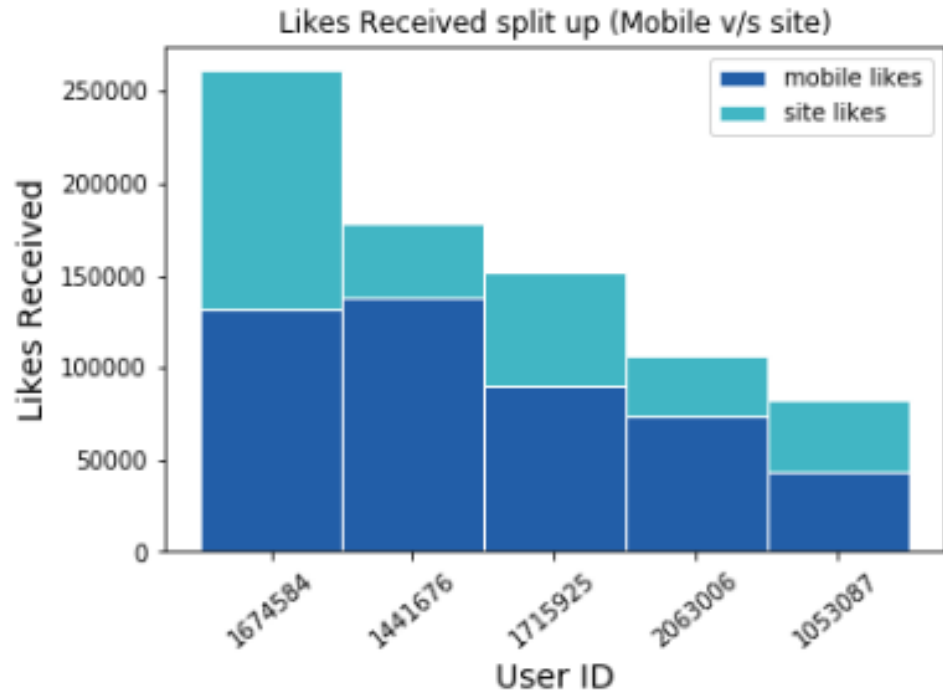
- No. of users in dataset depict a bump from 1940 to 1980 (50-60 year olds)
- Age Distribution shows peak between 15-28 years, then small peak between 45-55 years. It seems like ages of parents and kids of a generation.

Correlation Plot

- ▶ Friend count & friendships initiated are strongly correlated
- ▶ Ignore correlations of subsets of likes & likes received (mobile & site likes/likes received)
- ▶ Tenure and Age also seem to be related



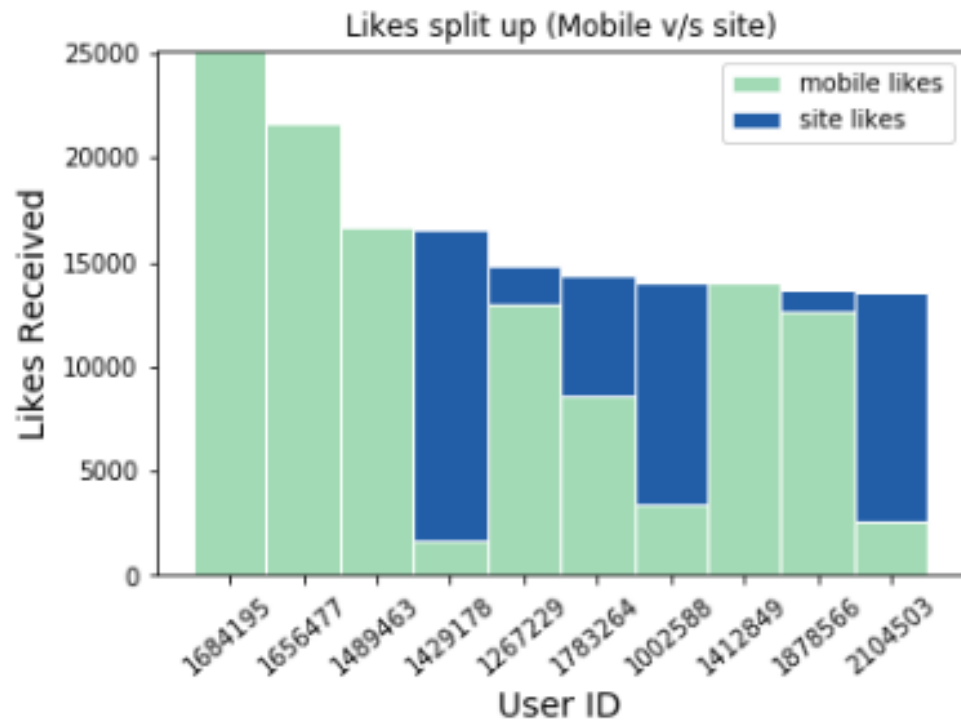
Likes / Likes Received



Likes Received

- ▶ It's a many to one relationship, many people give likes to a user.
- ▶ Top users likes show shift of people preferences towards mobile apps rather than site (mostly higher mobile likes than site likes)

Likes / Likes Received

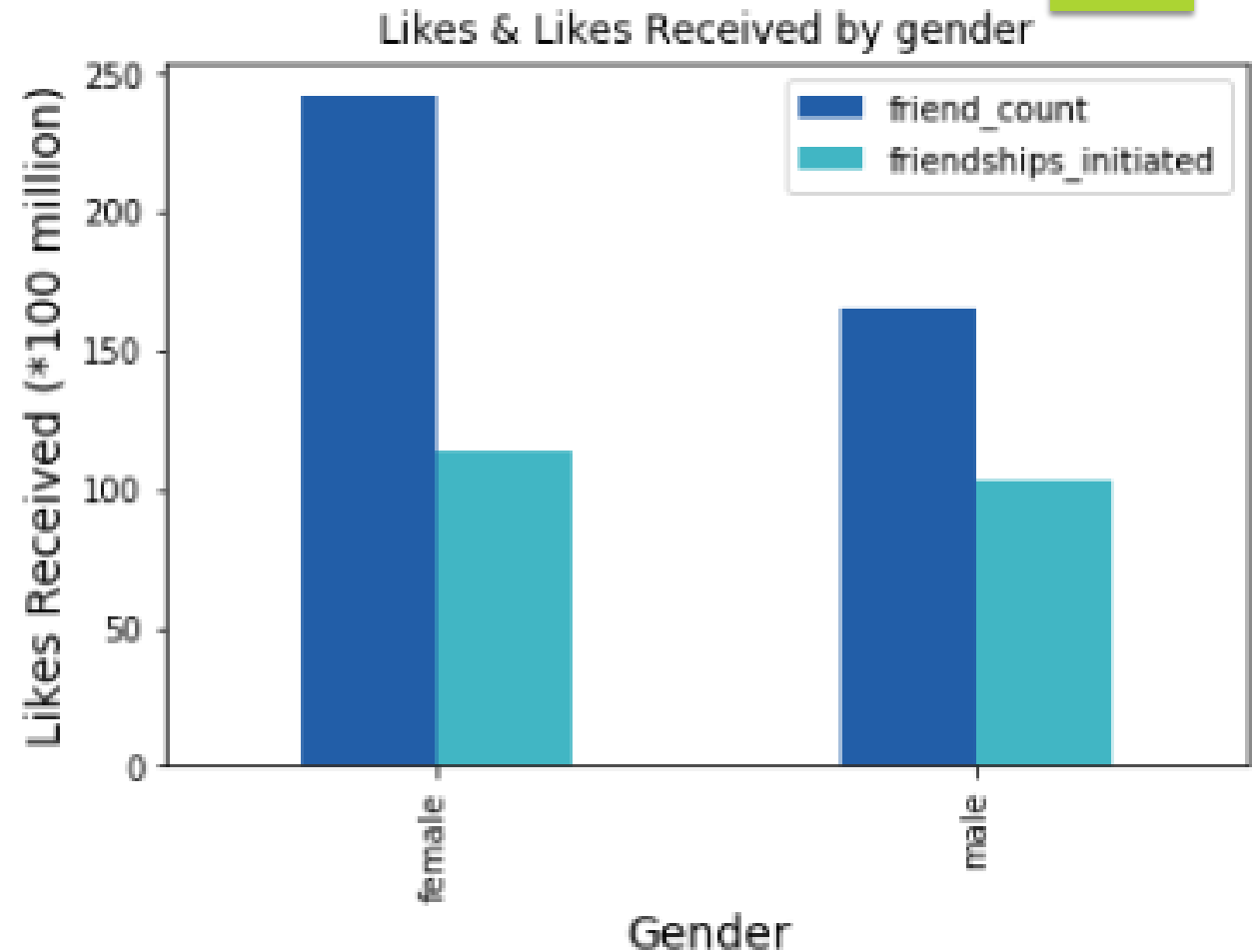


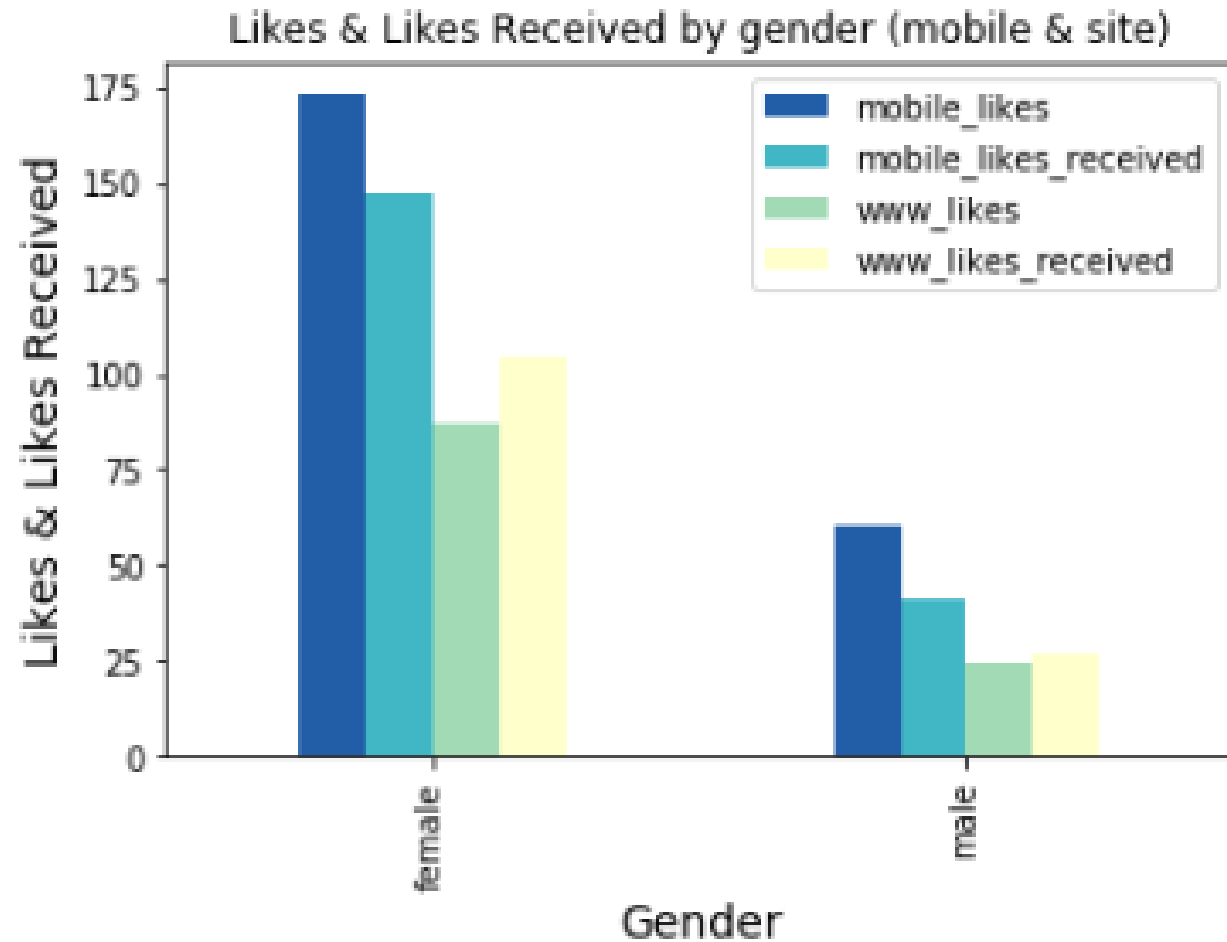
Likes given

- ▶ It's a one to many relationship, a user gives likes to many people.
- ▶ Top 4,6,7 & 10 user prefer the site more than mobile app
- ▶ Still a general skew towards mobile app

Gender Specific Interaction

- ▶ Women like as well as are liked a lot more than men (nearly 2.5 as much).
- ▶ Likes & likes received are comparable to each other for both men & women.
- ▶ Same behavior is true when the likes/like received are split up between mobile & site (nearly 3 times more for women than men)

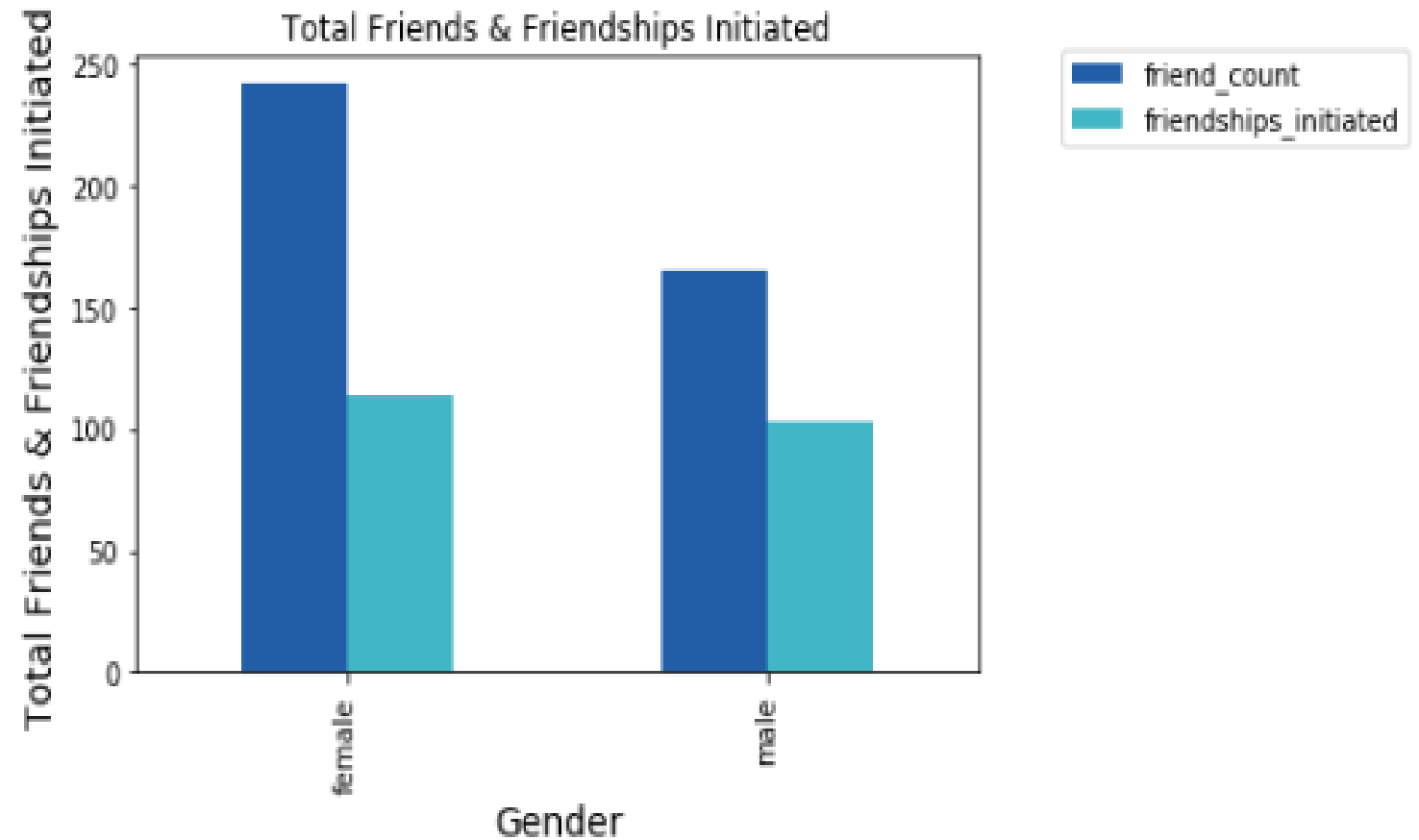




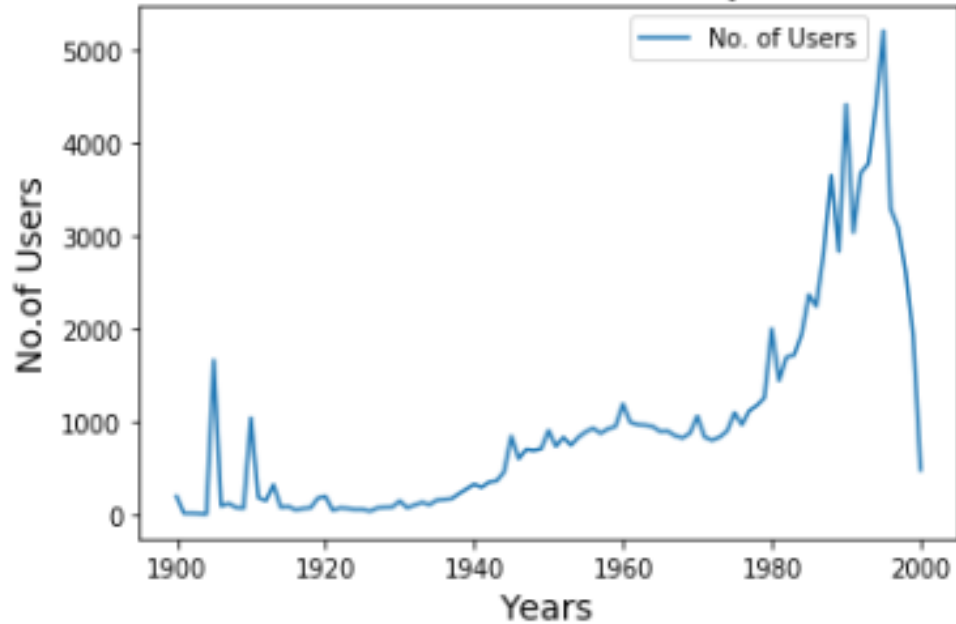
Average likes & likes received by men & women split up by medium

- Women receive and give nearly 3 time more likes both on mobile & site
- Likes received are more from site then likes given (raises question on mobile app popularity)

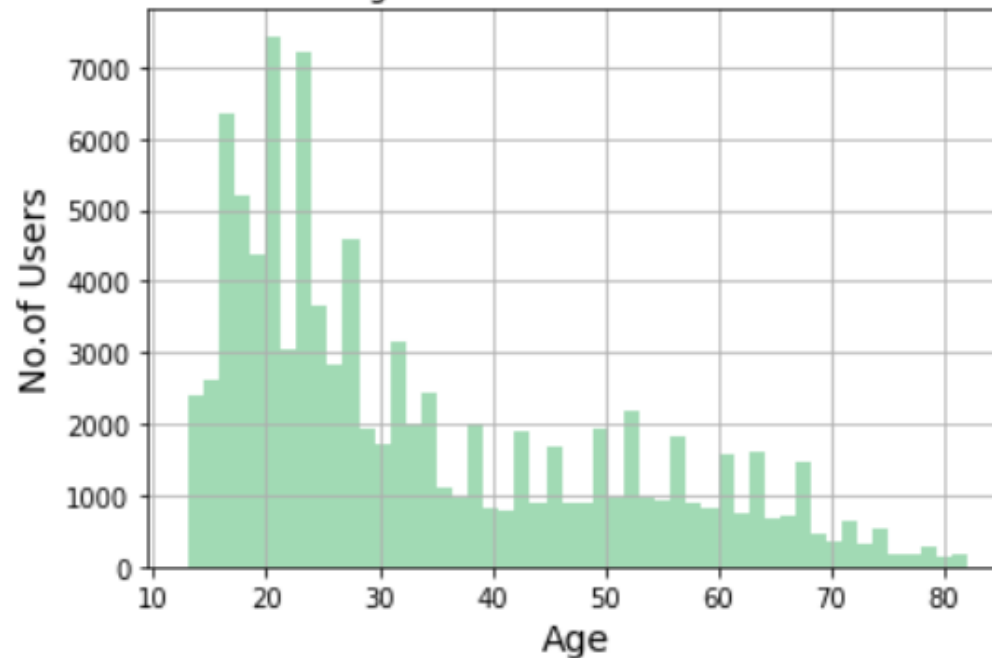
Women have more friends than men on facebook, the friendships initiated in proportion to friend count are more in case of men than women



Total no. of Users w.r.t birth year



Age wise distribution of Users



Users w.r.t birth year

► We see bumps between 1940 to 1980. After 1980 the no. users rocket. Since the data is till 2000 (we see miniscule value in 2000)

► Age Distribution shows peak between 15-28 years, then small peak between 45-55 years. It seems like ages of parents and kids of a generation.

Modelling

Equation: likes ~ age + tenure + friend_count + friendships_initiated + likes_received + gender

OLS Regression Results

Dep. Variable:	likes	R-squared:	0.199
Model:	OLS	Adj. R-squared:	0.199
Method:	Least Squares	F-statistic:	2867.
Date:	Mon, 03 Sep 2018	Prob (F-statistic):	0.00
Time:	20:57:10	Log-Likelihood:	-5.2665e+05
No. Observations:	69178	AIC:	1.053e+06
Df Residuals:	69171	BIC:	1.053e+06
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
C(gender)[female]	144.1977	4.661	30.935	0.000	135.061	153.334
C(gender)[male]	8.8286	4.057	2.176	0.030	0.878	16.779
Q("age")	-0.1658	0.094	-1.762	0.078	-0.350	0.019
Q("tenure")	0.0091	0.005	1.947	0.051	-5.95e-05	0.018
Q("friend_count")	0.1037	0.009	11.960	0.000	0.087	0.121
Q("friendships_initiated")	0.4793	0.018	27.312	0.000	0.445	0.514
Q("likes_received")	0.1388	0.002	84.549	0.000	0.136	0.142

