

A stylized illustration of Cupid, the Roman god of love, with golden curly hair, blue wings, and a quiver of arrows on his back. He is holding a bow and arrow, aiming towards the right. The illustration is set against a background of a globe with horizontal stripes, and the entire scene is framed by a decorative, wavy border in shades of pink and purple.

Online Dating

A Study of How to Get a Date

BUDT 733: Data Analysis for Decision Makers

May 11, 2006

Team 7

Vivek Arora
Tawney Bains
Ekta Mittal
Vikas Tiwari
Vishwas

Project Overview



The Client

An Online Dating Service Provider in Israel

Issue

Online Websites are facing more competition as communities become more fickle and unpredictable.

Websites need to attract larger audiences while keeping current users engaged.

Our Goal

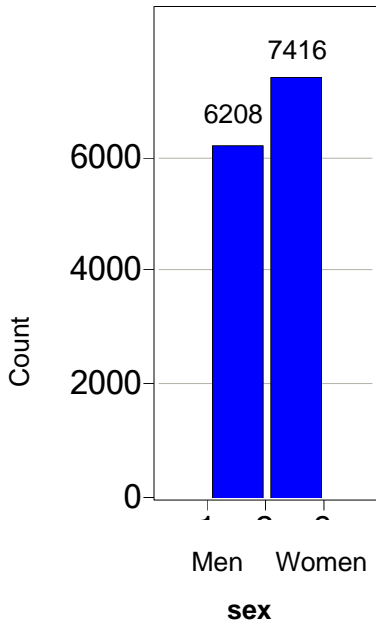
To explain and understand what contributes to a user's success in online dating

The Data

Users

45.6% Men

54.4% Women



- ♥ Profiles of over 13,000 online customers
- ♥ Each customer filled out unique fields – for example: sex, age, physical traits, job, behavioral traits and more!
- ♥ Online Photo Included or not
- ♥ Number of Messages Received (or not received)
- ♥ 42 Total Fields

Appearance

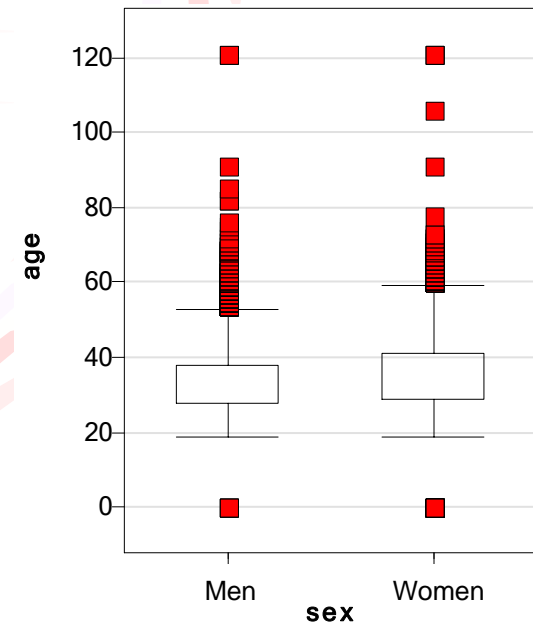
88.9% described their looks as Good or Great

Median Age Range

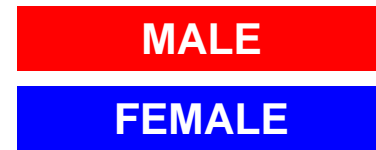
Overall: 28 to 40

Women: 29 to 41

Men: 28 to 38



Success Or Failure...



Definition:

Success: Message Received

Failure: No Message Received

Overall Results

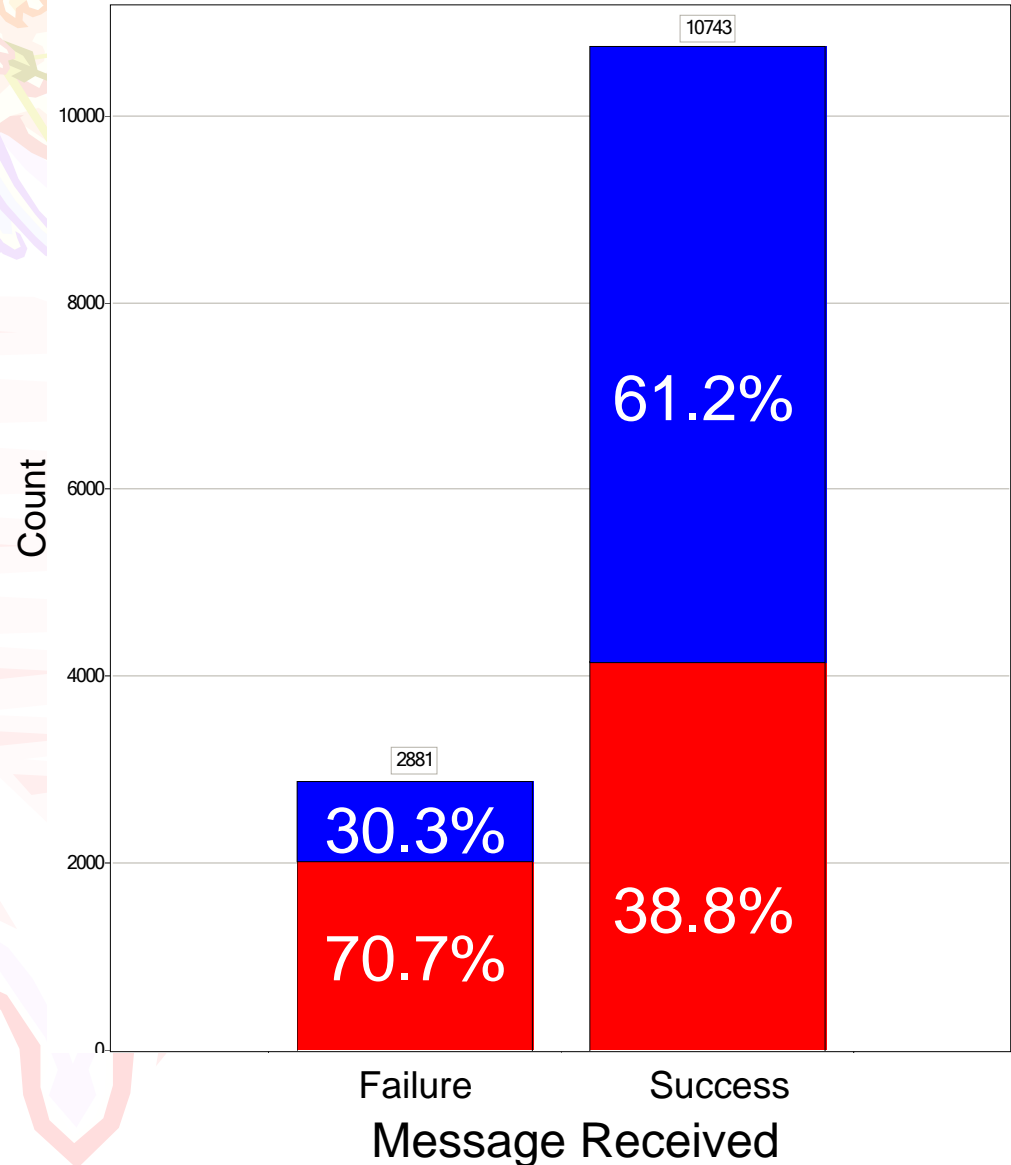
78.9% of users succeeded

Who's Receiving Messages?

Women!

Men are less successful

But there's hope...



Exploratory Analysis

Domain Knowledge:

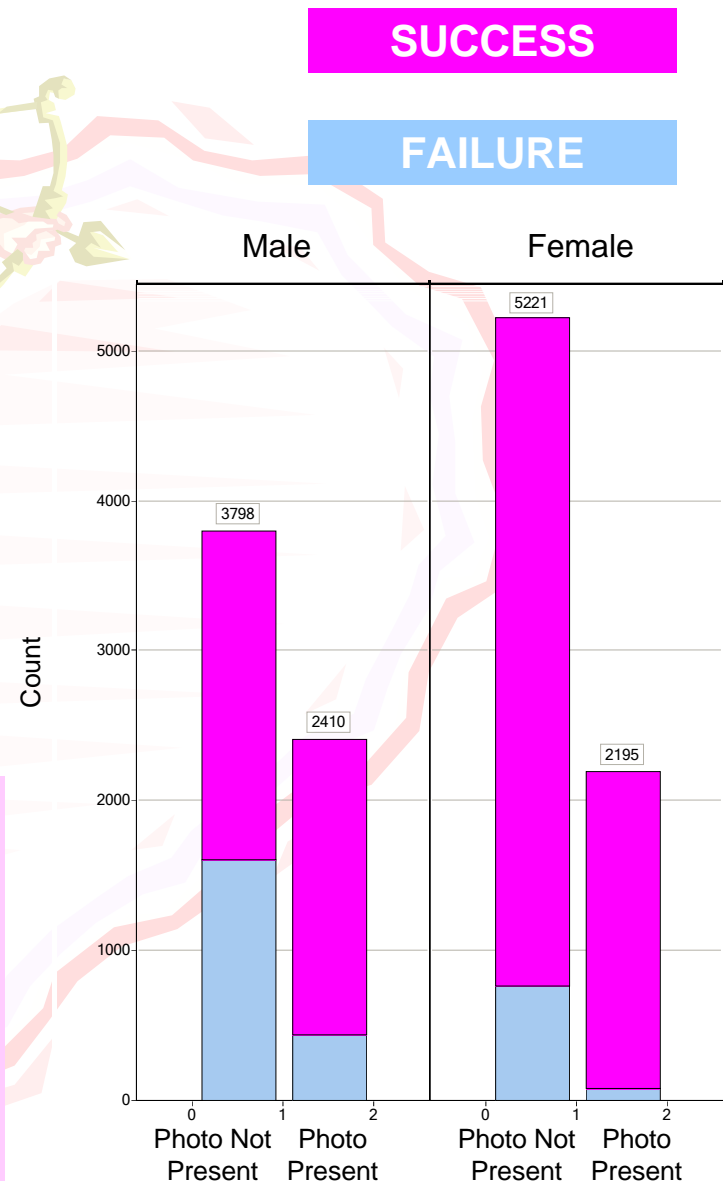
Posting a Photo is Important!

Failure: Overall Users

- ♥ Only 11.2% of those who posted a photo
- ♥ 26.2% of those who did not post a photo

Men: More important for you!

- ♥ 42.2% of all male users failed who did not post a photo
- ♥ Where as, only 18% of men failed with a photo



Exploratory Analysis

	Message Received		
Family_Status	Failure	Success	Grand Total
0	79.62%	20.38%	100.00%
1	21.51%	78.49%	100.00%
2	22.26%	77.74%	100.00%
3	13.31%	86.69%	100.00%
4	15.04%	84.96%	100.00%
5	26.83%	73.17%	100.00%
Grand Total	21.15%	78.85%	100.00%

Highly Correlated variables can be grouped together

Leaving a field blank, or skipping, leads to a much greater chance of failure than even putting a less “desirable” trait

	Message Received		
Appearance	Failure	Success	Grand Total
0	77.32%	22.68%	100.00%
1	17.40%	82.60%	100.00%
2	18.86%	81.14%	100.00%
3	33.33%	66.67%	100.00%
4	25.41%	74.59%	100.00%
5	20.20%	79.80%	100.00%
Grand Total	21.15%	78.85%	100.00%

Final Model Output

Input variables	Coefficient	Std. Error	p-value	Odds
Constant term	-3.99854207	0.43042201	0	*
Male	-1.79585683	0.07294627	0	0.16598517
age	0.02668369	0.00279212	0	1.02704287
is_gushdan/jeru/north	1.46395528	0.35838756	0.00004411	4.32302475
is_centre/sharon	1.59097469	0.36167061	0.00001088	4.90853071
is_south	1.29964232	0.3696737	0.00043868	3.66798425
family_status	-1.0990355	0.26959848	0.00004571	0.33319229
appearance	1.01995039	0.18021402	0.00000002	2.77305722
occupation	0.72185266	0.06786296	0	2.0582428
height	0.03822932	0.00790487	0.00000132	1.03896952
Personality_Traits	0.81053358	0.06555042	0	2.2491076
Requirements	1.13397706	0.09105827	0	3.10799265
photo_1	0.8766337	0.05883624	0	2.4027977
religion_background_1	0.77338248	0.1421386	0.00000005	2.16708398

Cut off Prob.Val. for Success (Updatable) **0.8**

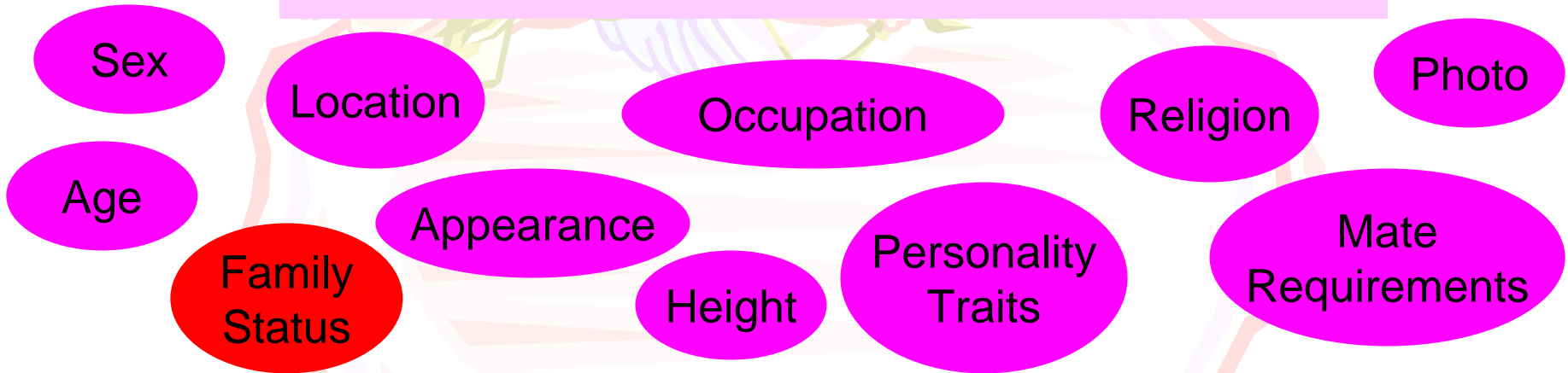
Classification Confusion Matrix		
	Predicted Class	
Actual Class	1	0
1	7939	2789
0	782	2067

Error Report			
Class	# Cases	# Errors	% Error
1	10728	2789	26.00
0	2849	782	27.45
Overall	13577	3571	26.30

Residual df	13563
Residual Dev.	10765.46484
% Success in training data	79.01598291
# Iterations used	11
Multiple R-squared	0.22829466

Insights & Conclusions

♥ Variables Impacting Success:



♥ Judea and Samaria: lower success rate than other locations

♥ People from the Center/Sharon area are the most successful

Recommendations



Combine Redundant Fields to encourage users to fill out everything increasing their chance of success



Make Fields that increase success “Required” or tell users that by filling in certain fields their chance of success increases



Make Fields that decrease success “Optional” For example fields that reveals one’s family status can be made optional



Encourage Men to Post Photos Online They will have greater success and will be more likely to stay with the site

Recommendations



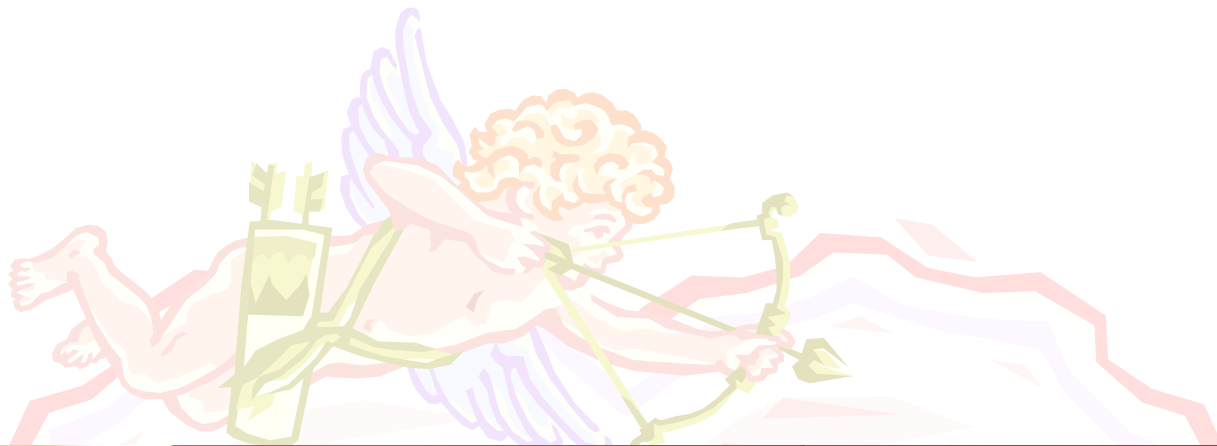
Offer Special Services for Women Promote that women users online have more success...men will follow when they know there are more women at a certain site



Use high regional success rate in marketing campaigns to attract more users from those regions



Test New Variables Other websites have fields like personal interests and dating status. This information may play an important role in user satisfaction and retention



05.06.2005 01:07

