



DS110 Final Project

# Divorce Prediction

Daniel Doh & Sophie Lim





# Problem Statement

- The goal of our project was to find out the most important factors that lead to divorce within married couples.

	0	1	2	3	4
1 If one of us apologizes when our discussion deteriorates, the discussion ends.	41%	5%	8%	28%	18%
2 I know we can ignore our differences, even if things get hard sometimes.	35%	14%	16%	22%	13%
3 When we need it, we can take our discussions with my spouse from the beginning and correct it.	30%	14%	15%	31%	10%
4 When I discuss with my spouse, to contact him will eventually work.	44%	7%	18%	19%	12%
5 The time I spent with my wife is special for us.	48%	6%	5%	26%	15%
6 We don't have time at home as partners.	51%	20%	17%	2%	1%
7 We are like two strangers who share the same environment at home rather than family.	67%	25%	3%	2%	3%
8 I enjoy our holidays with my wife.	48%	6%	12%	22%	12%
9 I enjoy traveling with my wife.	49%	4%	8%	29%	10%
10 Most of our goals are common to my spouse.	36%	11%	22%	20%	11%
11 I think that one day in the future, when I look back, I see that my spouse and I have been in harmony with each other.	42%	11%	2%	26%	19%
12 My spouse and I have similar values in terms of personal freedom.	34%	16%	13%	25%	12%
13 My spouse and I have similar sense of entertainment.	28%	19%	11%	28%	16%
14 Most of our goals for people (children, friends, etc) are the same.	39%	14%	12%	22%	13%
15 Our dreams with my spouse are similar and harmonious.	41%	12%	8%	30%	10%
16 We're compatible with my spouse about what love should be.	44%	8%	15%	21%	12%
17 We share the same views about being happy in our life with my spouse.	42%	9%	24%	23%	15%
18 My spouse and I have similar ideas about how marriage should be.	46%	6%	9%	20%	12%
19 My spouse and I have similar ideas about how roles should be in marriage.	45%	6%	3%	29%	16%
20 My spouse and I have similar values in trust.	48%	6%	12%	21%	13%
21 I know exactly what my wife likes.	46%	10%	10%	28%	6%
22 I know how my spouse wants to be taken care of when she's sick.	51%	8%	15%	17%	9%
23 I know my spouse's favorite food.	53%	4%	5%	25%	13%
24 I can tell you what kind of stress my spouse is facing in her/his life.	42%	9%	15%	21%	12%
25 I have knowledge of my spouse's inner world.	37%	16%	7%	25%	14%
26 I know my spouse's basic anxieties.	42%	12%	12%	22%	12%
27 I know what my spouse's current sources of stress are.	45%	11%	10%	27%	7%
28 I know my spouse's hopes and wishes.	50%	6%	18%	16%	10%
29 I know my spouse very well.	48%	8%	6%	25%	14%
30 I know my spouse's friends and their social relationships.	42%	11%	14%	21%	12%
31 I feel aggressive when I argue with my spouse.	26%	16%	12%	10%	35%

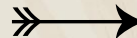
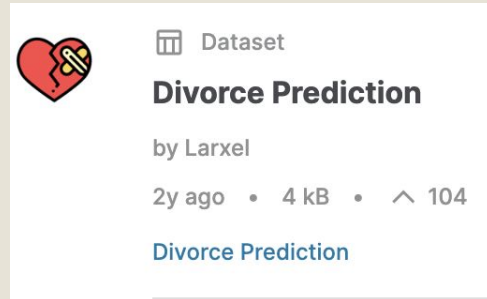
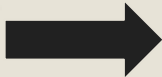
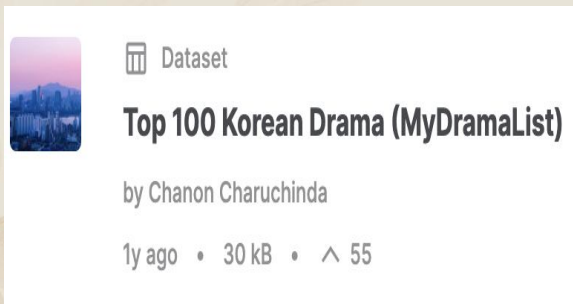
32 When discussing with my spouse, I usually use expressions such as 'you always' or 'you never'.	27%	18%	7%	19%	29%
33 I can use negative statements about my spouse's personality during our discussions.	42%	12%	4%	10%	33%
34 I can use offensive expressions during our discussions.	29%	22%	6%	15%	28%
35 I can insult my spouse during our discussions.	50%	8%	1%	9%	33%
36 I can be humiliating when we discuss.	52%	5%	1%	14%	28%
37 My discussion with my spouse is not calm.	29%	18%	6%	9%	38%
38 I hate my spouse's way of open a subject.	38%	14%	4%	14%	31%
39 Our discussions often occur suddenly.	29%	16%	8%	8%	38%
40 We're just starting a discussion before I know what's going on.	42%	8%	4%	14%	33%
41 When I talk to my spouse about something, my calm suddenly breaks.	32%	15%	8%	9%	35%
42 When I argue with my spouse, I only go out and I don't say a word.	26%	10%	15%	20%	29%
43 I mostly stay silent to calm the environment a little bit.	9%	11%	19%	19%	41%
44 Sometimes I think it's good for me to leave home for a while.	35%	9%	10%	16%	29%
45 I'd rather stay silent than discuss with my spouse.	16%	14%	12%	21%	36%
46 Even if I'm right in the discussion, I stay silent to hurt my spouse.	13%	11%	16%	28%	32%
47 When I discuss with my spouse, I stay silent because I am afraid of not being able to control my anger.	19%	20%	12%	12%	37%
48 I feel right in our discussions.	6%	6%	28%	29%	31%
49 I have nothing to do with what I've been accused of.	16%	16%	15%	16%	36%
50 I'm not actually the one who's guilty about what I'm accused of.	11%	18%	21%	15%	34%
51 I'm not the one who's wrong about problems at home.	7%	17%	25%	22%	28%
52 I wouldn't hesitate to tell my spouse about her/his inadequacy.	14%	16%	14%	18%	39%
53 When I discuss, I remind my spouse of her/his inadequacy.	18%	18%	15%	18%	31%
54 I'm not afraid to tell my spouse about her/his incompetence.	29%	16%	12%	9%	34%





# Backstory

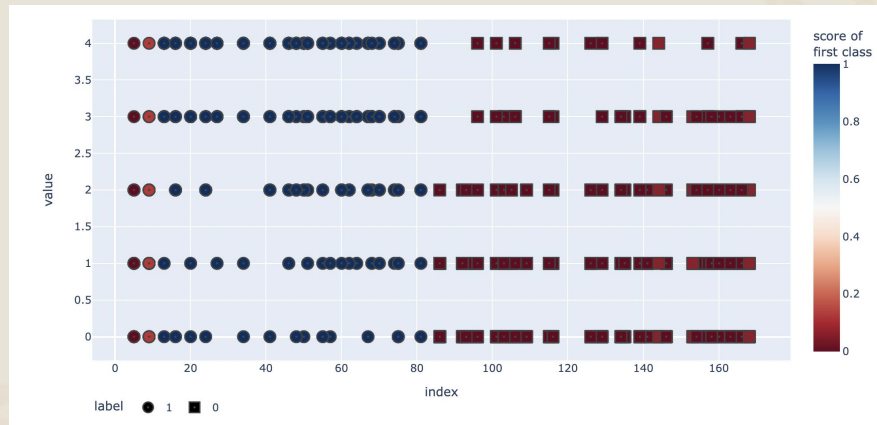
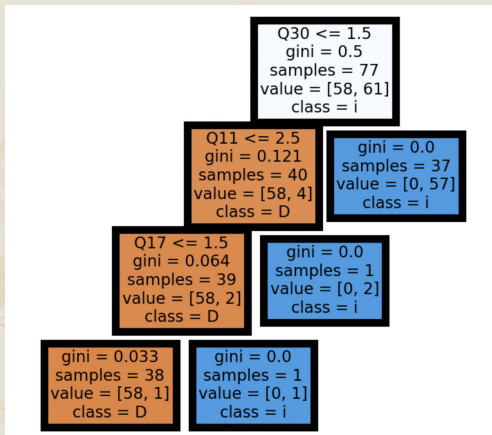
- Originally, we wanted to analyze a dataset that was interesting to us (Top 100 kdramas). However, we found that the dataset contained no trends for analysis, so we switched to another data set containing factors predicting divorce.





# Methods

- We used three different classification models: Logistic Regression, K-Nearest Neighbors, and Random Forest.





# Challenges

- With our new data set, our data was very categorical, so we couldn't initially get nice visualizations. We overcame this issue by using Python visualization libraries such as Shap and plotly express.

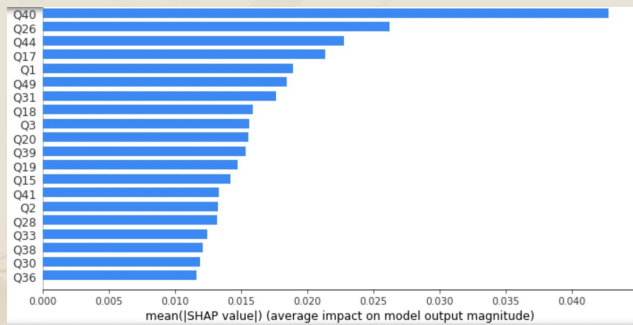




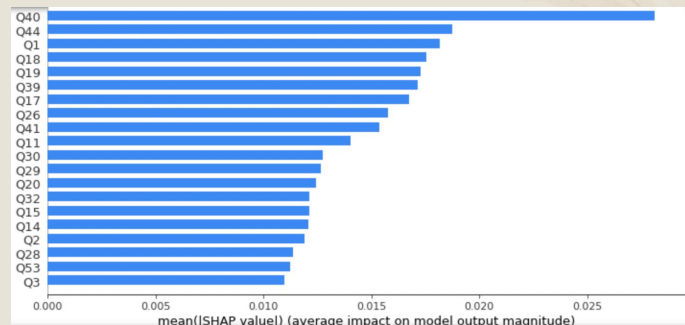
# Results

- Our models differed slightly in which factors they weighted more in their analysis. But they all tended to agree that measurements of aggression, communication, and goals influenced divorce the most.

## Logistic Regression/ Random Forest



## K- Nearest Neighbors





# Conclusions

- We do have suspicions of overfitting since our models scored exceptionally well. However, that's not something we can check unless we get more data.
- We learned how to use scikit-learn modules and other machine learning visualization libraries such as Shap and plotly express.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	26
1	1.00	1.00	1.00	25
accuracy			1.00	51
macro avg	1.00	1.00	1.00	51
weighted avg	1.00	1.00	1.00	51

```
[[26  0]
 [ 0 25]]
```







# Fun Fact

- Participants who claimed to know their spouse's friends and social relationships very well had a higher divorce rate than those who claimed they didn't.

