Pune 2018

Classification with the Somerville Dataset

John Stinson - johnstinson99@gmail.com

Arlington
60
3
Tufts University
DAVIS SQUARE
Somerville
28
93
99
Bunker Hill Monu
Harvard University
Mt Auburn Cemetery
3
Cambridge
16
TD Garden
Binney St
The
Harvard Business School
Soldiers Field Rd
2A
Boston
Satellite
ALLSTON
Boston Children's Museum
Google
BACK BAY
30

https://catalog.data.gov/dataset?tags=happiness

DATA.GOV

DATA TOPICS IMPACT APPLICATIONS DEVELOPERS CONTACT

DATA CATALOG

/ Datasets Organizations

Search datasets...

Order by: Popular

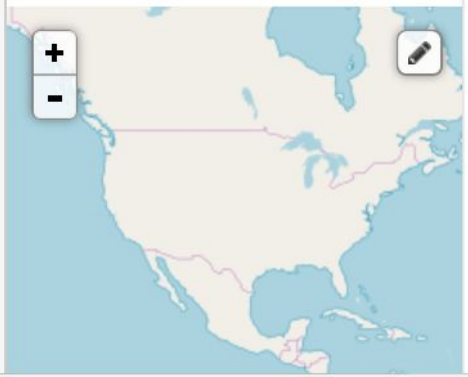
Datasets ordered by Popular

Tags:

happiness ✖

Filter by location Clear

Enter location...



2 datasets found

Somerville Happiness Survey responses - 2011, 2013, 2015

21 recent views

City of Somerville — The City of Somerville sends out a happiness survey to a random sample of Somerville residents asking them to rate their personal happiness and their satisfaction...

CSV RDF JSON XML

City

2013 Somerville Resident Happiness Survey

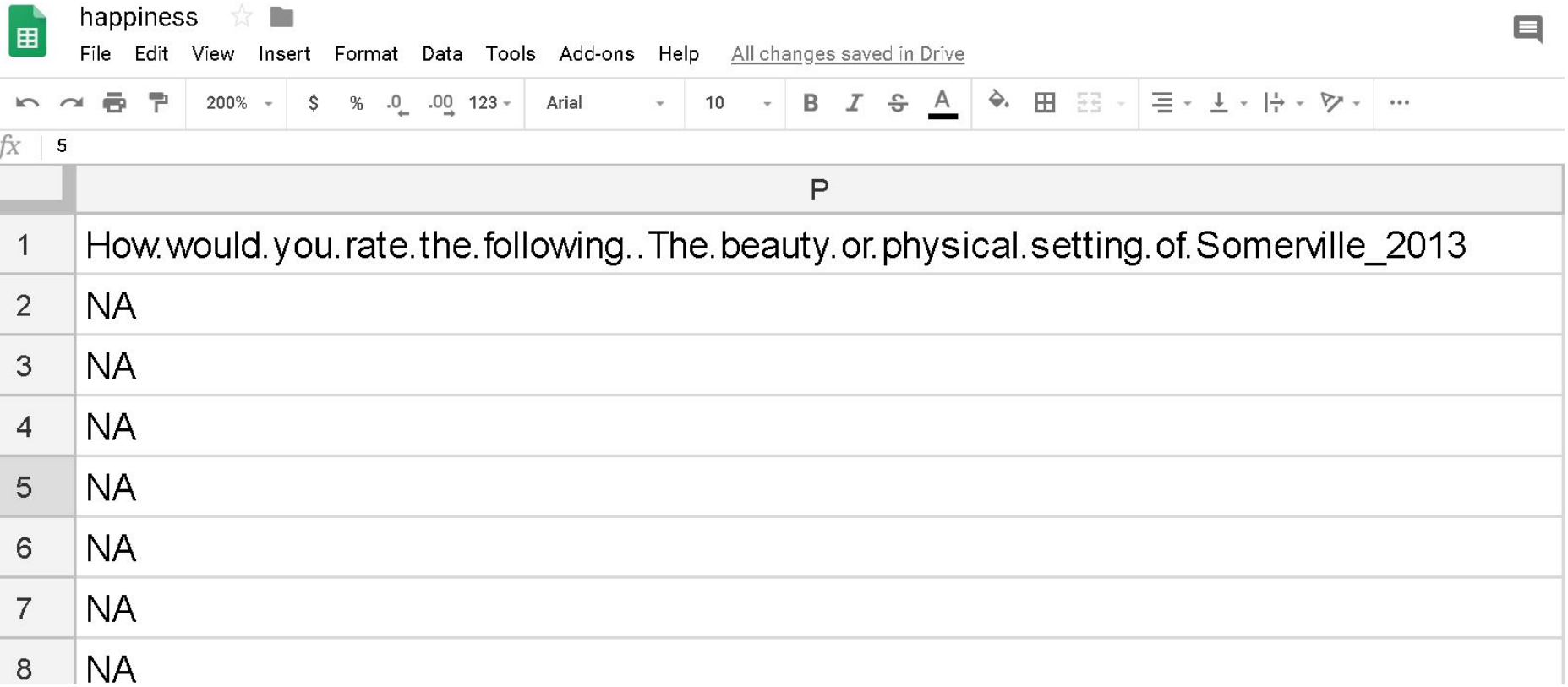
City of Somerville — 2013 Somerville Resident Happiness Survey

City

Activate Wind
Go to Settings to a

B	C	D
Year	How.happy.do.you.feel.right.now.	How.satisfied.are.you.with.your.life.in.general.
2015	7	5
2015	8	8
2015	5	9
2015	5	8
2015	7	9
2015	7	8
2015	10	10
2015	7	8

Demo


happiness
File Edit View Insert Format Data Tools Add-ons Help
All changes saved in Drive
200% $ % .0 .00 123 Arial 10
fx 5
P
1 How.would.you.rate.the.following..The.beauty.or.physical.setting.of.Somerville_2013
2 NA
3 NA
4 NA
5 NA
6 NA
7 NA
8 NA

happiness

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

200% $ % .0 .00 123 Arial 10 B I S A

fx R

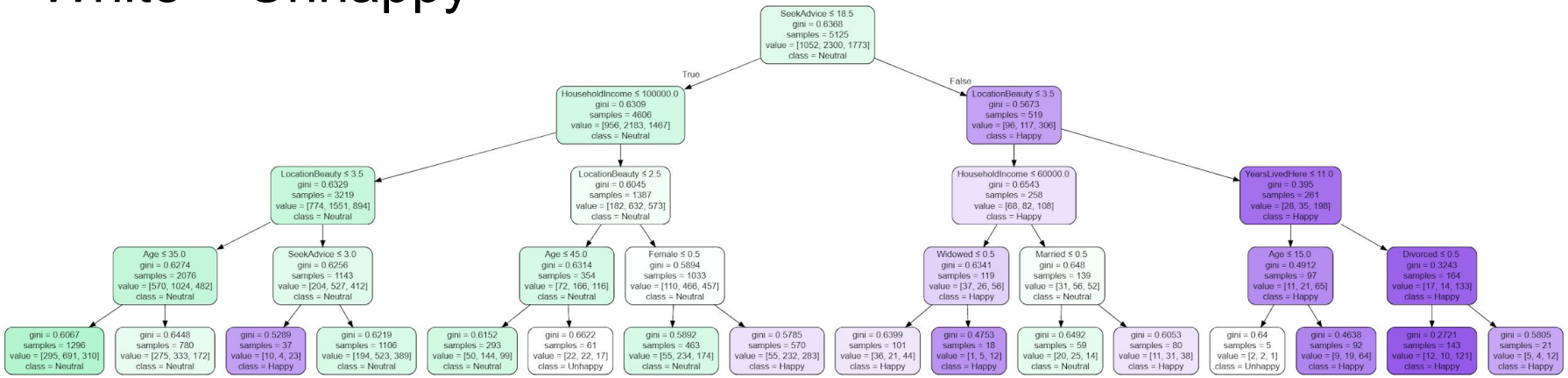
	B	C	
122	2011	i	
123	2011	9*	
124	2011	9*	
125	2011	810	
126	2011	88	
127	2011	77	
128	2011	57	
129	2011	42	
130	2015	10	

Demo: Data Munging with Pandas:
Somerville Happiness Dataset

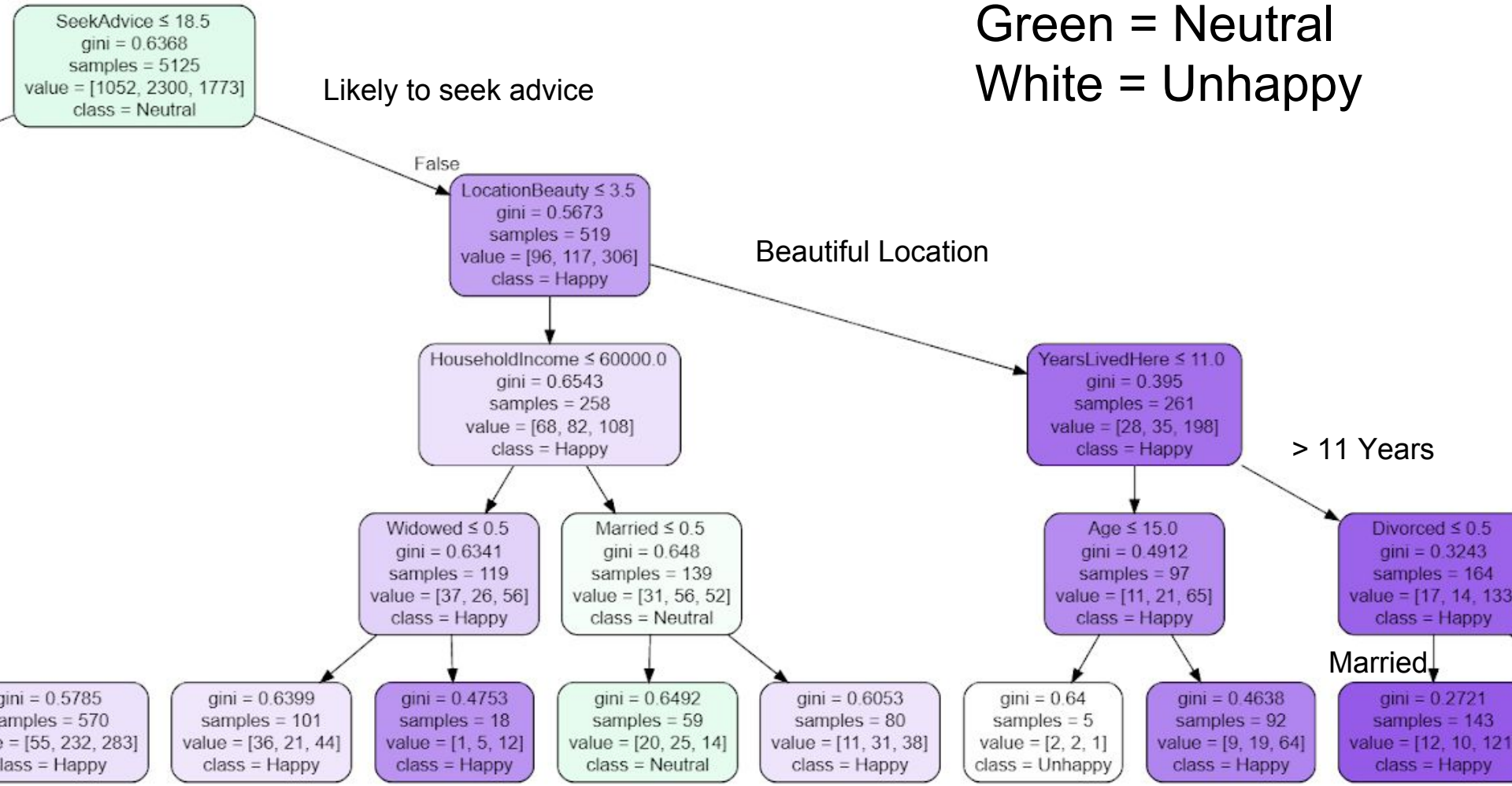
- Different columns each year.
- Free text fields,so invalid characters
- Many irrelevant columns
- Long column names
- Strings in numeric columns eg '0-5 years'
- Non-sequential values – e.g. Marital Status

Decision Tree - 5 levels

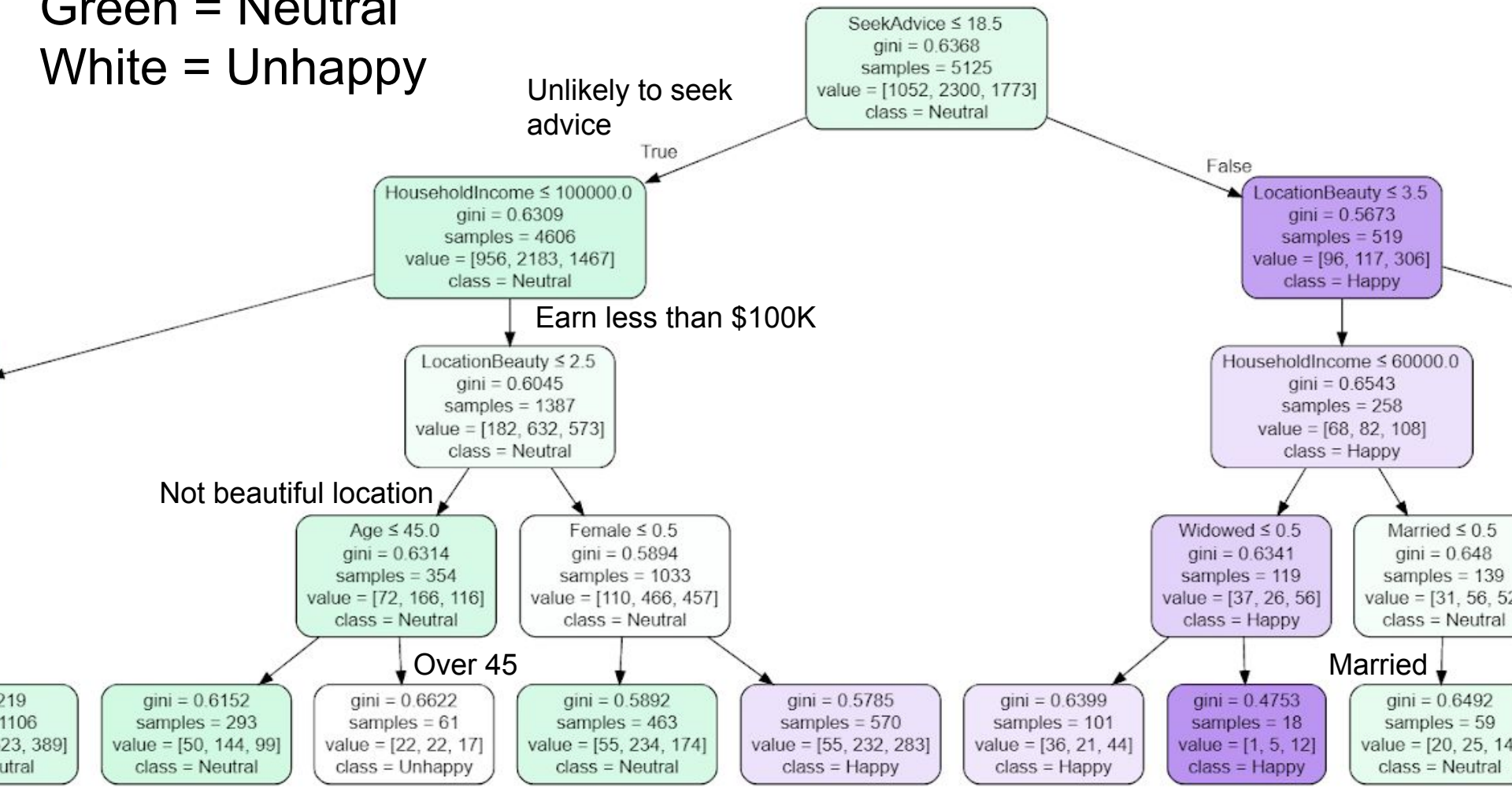
Key:
Purple = Happy
Green = Neutral
White = Unhappy



Key:
Purple = Happy
Green = Neutral
White = Unhappy



Key:
Purple = Happy
Green = Neutral
White = Unhappy



Next Steps

Experiment with

- Different number of levels of decision tree
- Including/additional variables
- Use alternative output columns
- Alternative classification models

Questions?

John Stinson - johnstinson99@gmail.com