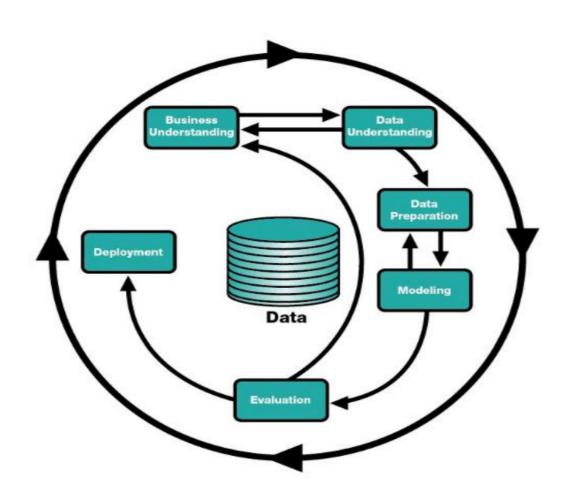


Methodology





- 1. Business understanding
- 2. Data understanding
- 3. Data preparation
- 4. Modelling
- 5. Evaluation
- 6. Recommendation

Baseline Factors for Quality of life in the US

Baseline Factors

- Poor Physical Health days
- Median household income
- Unemployment rate
- R-squared = 0.851

All our coefficients are significant. So we recommend taking these variables into consideration when choosing where to immigrate to, in the US.

Dep. Variable:	Poor_mental_health_days_raw_value	R-squared:	0.851
Model:	OLS	Adj. R-squared:	0.851
Method:	Least Squares	F-statistic:	5965.
Date:	Wed, 26 Feb 2020	Prob (F-statistic):	0.00
Time:	14:05:37	Log-Likelihood:	65.095
No. Observations:	3142	AIC:	-122.2
Df Residuals:	3138	BIC:	-97.98
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	0.5142	0.048	10.820	0.000	0.421	0.607
Poor_physical_health_days_raw_value	0.8521	0.009	99.116	0.000	0.835	0.869
Median_household_income_raw_value	3.125e-06	4.15e-07	7.529	0.000	2.31e-06	3.94e-06
Unemployment_raw_value	-1.8160	0.303	-5.991	0.000	-2.410	-1.222

on: (Durbin-Watson:	5.801	Omnibus:
3): 6	Jarque-Bera (JB):	0.055	Prob(Omnibus):
3): 0.0	Prob(JB):	0.005	Skew:
lo. 3.79	Cond. No.	3.226	Kurtosis:

Factors for Quality of Life in the US



Theoretically,We want a county with:

- * Low Frequent_mental_distress
- * Low Poor_physical_health_days
- * Low Frequent_physical_distress
- * Low Poor_or_fair_health
- * Low Adult_smoking
- * Sufficient_sleep
- * Low Diabetes_prevalence
- * Low Children_in_poverty
- * High Life_expectancy
- * Low Excessive_drinking
- * Low Food_insecurity



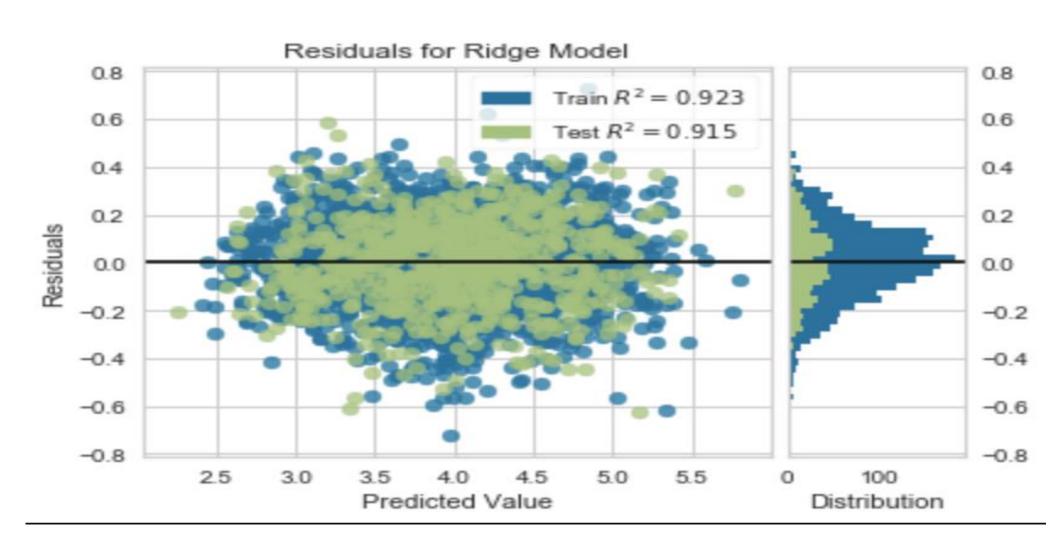




- * Low Premature_death
- * Medium/high Median_household_income
- * Low Physical_inactivity
- * Low Teen_births_rate
- * A good_employment_rate
- * Low Drug_overdose_deaths
- * Good Income_equality

Ridge / Lasso / Elastic Net Model





Ridge Model

~	

variables	weights
Insufficient_sleep_raw_value^2	-0.344677
cient_sleep_raw_value Food_insecurity_r	-0.320856
sical_health_days_raw_value Excessive	-0.312443
_fair_health_raw_value Life_expectancy	-0.273512
Frequent_physical_distress_raw_value^2	-0.258313
•••	•••
_fair_health_raw_value Adult_smoking_ra	0.266736
nt_sleep_raw_value Unemployment_raw	0.340726
cient_sleep_raw_value Excessive_drinkin	0.347430
sical_health_days_raw_value Life_expec	0.417395
Poor_physical_health_days_raw_value	0.460157

Future work slide



- Run survey to determine endogenous driver
- Branch into new industry with strongly predictive model of Poor Mental Health Days per Month
- Run an experiment to confirm whether sleep deprivation increases life satisfaction.



THANKYOU Q&A