Team Members: Matthew Kirdin, Vladislav Pendishchuk

Chosen Dataset: <u>Diamonds Price Dataset</u>

Reason of Choice: We choose Diamond Price Dataset because it has a variety of easily interpretable features as well as an abundance of data.

Dataset Info: This dataset contains 50,000 entries and 10 features (listed below).

Feature	Туре	Description, units of measurement
carat	quantitative	Weight, carats
cut	categorical	Cutting quality
color	categorical	Color from J (worst) to D (best)
clarity	categorical	Clarity from left to right is worst to best: 11, S12, S11,
		VS2, VS1, VVS2, VVS1, IF
X	quantitative	Length, mm
У	quantitative	Width, mm
Z	quantitative	
depth	quantitative	Percentage depth $\frac{z}{mean(x,y)}$, %
table	quantitative	Width of the widest point at the top of the
		diamond, mm
price	quantitative	Diamond price

Data examples:

	carat	cut	color	clarity	depth	table	price	x	У	z
0	0.23	Ideal	Е	SI2	61.5	55.0	326	3.95	3.98	2.43
1	0.21	Premium	Е	SI1	59.8	61.0	326	3.89	3.84	2.31
2	0.23	Good	Е	VS1	56.9	65.0	327	4.05	4.07	2.31
3	0.29	Premium	1	VS2	62.4	58.0	334	4.20	4.23	2.63
4	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75