Samuel Wallace
Connor McInnes
ECE102 Project Report

Gas prices fluctuate in a frustratingly unpredictable manner. Anyone who does not realize this likely does not need to purchase it often. It was with this frustration in mind that we set about trying to solve an often found problem: There is an enormous amount of data out there on the web, but most of the time it is not held in a format which is able to be easily imported into tools where that data can actually be put to use. The applications of this data are left to whatever analysis was included by the authors who published the data originally. Attempting further analysis of data available on the web can be tedious and time consuming at best. For our final project, we set about showing a method of parsing data from an online source into a workable format– in this case, gas price data from the United States Department of Labor: Bureau of Labor Statistics. The data provided by the Bureau of Labor Statistics is extensive, listing the average price per gallon of gasoline every month of every year dating back to 1976 (seen in Figure 1) but this data is stuck on an html webpage. It cannot be analyzed, or even simply displayed in more visually appealing way.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1976	0.605	0.600	0.594	0.592	0.600	0.616	0.623	0.628	0.630	0.629	0.629	0.626
1977	0.627	0.637	0.643	0.651	0.659	0.665	0.667	0.667	0.666	0.665	0.664	0.665
1978	0.648	0.647	0.647	0.649	0.655	0.663	0.674	0.682	0.688	0.690	0.695	0.705
1979	0.716	0.730	0.755	0.802	0.844	0.901	0.949	0.988	1.020	1.028	1.041	1.065
1980	1.131	1.207	1.252	1.264	1.266	1.269	1.271	1.267	1.257	1.250	1.250	1.258
1981	1.298	1.382	1.417	1.412	1.400	1.391	1.382	1.376	1.376	1.371	1.369	1.365
1982	1.358	1.334	1.284	1.225	1.237	1.309	1.331	1.323	1.307	1.295	1.283	1.260
1983	1.230	1.187	1.152	1.215	1.259	1.277	1.288	1.285	1.274	1.255	1.241	1.231
1984	1.216	1.209	1.210	1.227	1.236	1.229	1.212	1.196	1.203	1.209	1.207	1.193
1985	1.148	1.131	1.159	1.205	1.231	1.241	1.242	1.229	1.216	1.204	1.207	1.208
1986	1.194	1.120	0.981	0.888	0.923	0.955	0.890	0.843	0.860	0.831	0.821	0.823
1987	0.862	0.905	0.912	0.934	0.941	0.958	0.971	0.995	0.990	0.976	0.976	0.961
1988	0.933	0.913	0.904	0.930	0.955	0.955	0.967	0.987	0.974	0.957	0.949	0.930
1989	0.918	0.926	0.940	1.065	1.119	1.114	1.092	1.057	1.029	1.027	0.999	0.980
1990	1.042	1.037	1.023	1.044	1.061	1.088	1.084	1.190	1.294	1.378	1.377	1.354
1991	1.247	1.143	1.082	1.104	1.156	1.160	1.127	1.140	1.143	1.122	1.134	1.123
1992	1.073	1.054	1.058	1.079	1.136	1.179	1.174	1.158	1.158	1.154	1.159	1.136
1993	1.117	1.108	1.098	1.112	1.129	1.130	1.109	1.097	1.085	1.127	1.113	1.070
1994	1.043	1.051	1.045	1.064	1.080	1.106	1.136	1.182	1.177	1.152	1.163	1.143
1995	1.129	1.120	1.115	1.140	1.200	1.226	1.195	1.164	1.148	1.127	1.101	1.101
1996	1.129	1.124	1.162	1.251	1.323	1.299	1.272	1.240	1.234	1.227	1.250	1.260
1997	1.261	1.255	1.235	1.231	1.226	1.229	1.205	1.253	1.277	1.242	1.213	1.177
1998	1.131	1.082	1.041	1.052	1.092	1.094	1.079	1.052	1.033	1.042	1.028	0.986
1999	0.972	0.955	0.991	1.177	1.178	1.148	1.189	1.255	1.280	1.274	1.264	1.298
2000	1.301	1.369	1.541	1.506	1.498	1.617	1.593	1.510	1.582	1.559	1.555	1.489
2001	1.472	1.484	1.447	1.564	1.729	1.640	1.482	1.427	1.531	1.362	1.263	1.131
2002	1.139	1.130	1.241	1.407	1.421	1.404	1.412	1.423	1.422	1.449	1.448	1.394
2003	1.473	1.641	1.748	1.659	1.542	1.514	1.524	1.628	1.728	1.603	1.535	1.494
2004	1.592	1.672	1.766	1.833	2.009	2.041	1.939	1.898	1.891	2.029	2.010	1.882
2005	1.823	1.918	2.065	2.283	2.216	2.176	2.316	2.506	2.927	2.785	2.343	2.186
2006	2.315	2.310	2.401	2.757	2.947	2.917	2.999	2.985	2.589	2.272	2.241	2.334
2007	2.274	2.285	2.592	2.860	3.130	3.052	2.961	2.782	2.789	2.793	3.069	3.020
2008	3.047	3.033	3.258	3.441	3.764	4.065	4.090	3.786	3.698	3.173	2.151	1.689
2009	1.787	1.928	1.949	2.056	2.265	2.631	2.543	2.627	2.574	2.561	2.660	2.621
2010	2.731	2.659	2.780	2.858	2.869	2.736	2.736	2.745	2.704	2.795	2.852	2.985
2011	3.091	3.167	3.546	3.816	3.933	3.702	3.654	3.630	3.612	3.468	3.423	3.278
2012	3.399	3.572	3.868	3.927	3.792	3.552	3.451	3.707	3.856	3.786	3.488	3.331
2013	3.351	3.693	3.735	3.590	3.623	3.633	3.628	3.600	3.556	3.375	3.251	3.277
2014	3.320	3.364	3.532	3.659	3.691	3.695	3.633	3.481	3.403	3.182	2.887	2.560
2015	2.110	2.249	2.483	2.485	2.775	2.832	2.832	2.679	2.394	2.289	2.185	2.060
2016	1.967	1.767	1.958									

(Figure 1)

The first step of our project was to download a local copy of the html source of this page to analyze with a program we would write. We had initially planned to access the webpage remotely through our program, but displaying the full set of data we needed required some initial interaction with some JavaScript elements on the webpage. Programmatically interacting with JavaScript elements to display the information we wanted on the page would have been possible, but was beyond the scope of this project and, quite frankly, could have been an entire

The entirety of this project (including source, presentation, and this report) can be found on GitHub: Here

project on its own. From here we went about writing a program to extract the information we wanted from the webpage and storing it in a more convenient format. The basic logic of the program will be explained in this report, but we won't go into extreme technical detail as the (well commented) source code of the program is linked in the git repository referenced in the footer.

The program essentially functioned by copying the entire source file to a two-dimensional array and then trimming this array further and further until only the relevant data remained. Step one was opening up the html source file and writing each line to a consecutive row of a two-dimensional character array. This array was then used to create a new character array containing only the lines with relevant data on them. This new array was then iterated over in a for loop utilizing some ASCII value and if/else trickery to eliminate unnecessary characters leaving only relevant data points and white space in yet another two dimensional array. This final 'work-in-progress' array was then written to a text file which was itself read into the program again by the fscanf function (which conveniently ignores whitespace) creating a final 2D array which contained only the relevant data points.

This final array was then utilized in the manual creation of a basic excel spreadsheet in an xml format. With this initial excel sheet (show below) in hand, it was quite easy to convert this basic spreadsheet (Figure 2) into a much more detailed and visually appealing one (Figure 3).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	1976	0.605	0.6	0.594	0.592	0.6	0.616	0.623	0.628	0.63	0.629	0.629	0.626
2	1977	0.627	0.637	0.643	0.651	0.659	0.665	0.667	0.667	0.666	0.665	0.664	0.665
3	1978	0.648	0.647	0.647	0.649	0.655	0.663	0.674	0.682	0.688	0.69	0.695	0.705
4	1979	0.716	0.73	0.755	0.802	0.844	0.901	0.949	0.988	1.02	1.028	1.041	1.065
5	1980	1.131	1.207	1.252	1.264	1.266	1.269	1.271	1.267	1.257	1.25	1.25	1.258
6	1981	1.298	1.382	1.417	1.412	1.4	1.391	1.382	1.376	1.376	1.371	1.369	1.365
7	1982	1.358	1.334	1.284	1.225	1.237	1.309	1.331	1.323	1.307	1.295	1.283	1.26
8	1983	1.23	1.187	1.152	1.215	1.259	1.277	1.288	1.285	1.274	1.255	1.241	1.231
9	1984	1.216	1.209	1.21	1.227	1.236	1.229	1.212	1.196	1.203	1.209	1.207	1.193
10	1985	1.148	1.131	1.159	1.205	1.231	1.241	1.242	1.229	1.216	1.204	1.207	1.208
11	1986	1.194	1.12	0.981	0.888	0.923	0.955	0.89	0.843	0.86	0.831	0.821	0.823
12	1987	0.862	0.905	0.912	0.934	0.941	0.958	0.971	0.995	0.99	0.976	0.976	0.961
13	1988	0.933	0.913	0.904	0.93	0.955	0.955	0.967	0.987	0.974	0.957	0.949	0.93
14	1989	0.918	0.926	0.94	1.065	1.119	1.114	1.092	1.057	1.029	1.027	0.999	0.98
15	1990	1.042	1.037	1.023	1.044	1.061	1.088	1.084	1.19	1.294	1.378	1.377	1.354
16	1991	1.247	1.143	1.082	1.104	1.156	1.16	1.127	1.14	1.143	1.122	1.134	1.123
17	1992	1.073	1.054	1.058	1.079	1.136	1.179	1.174	1.158	1.158	1.154	1.159	1.136
18	1993	1.117	1.108	1.098	1.112	1.129	1.13	1.109	1.097	1.085	1.127	1.113	1.07
19	1994	1.043	1.051	1.045	1.064	1.08	1.106	1.136	1.182	1.177	1.152	1.163	1.143
20	1995	1.129	1.12	1.115	1.14	1.2	1.226	1.195	1.164	1.148	1.127	1.101	1.101
21	1996	1.129	1.124	1.162	1.251	1.323	1.299	1.272	1.24	1.234	1.227	1.25	1.26
22	1997	1.261	1.255	1.235	1.231	1.226	1.229	1.205	1.253	1.277	1.242	1.213	1.177
23	1998	1.131	1.082	1.041	1.052	1.092	1.094	1.079	1.052	1.033	1.042	1.028	0.986
24	1999	0.972	0.955	0.991	1.177	1.178	1.148	1.189	1.255	1.28	1.274	1.264	1.298
25	2000	1.301	1.369	1.541	1.506	1.498	1.617	1.593	1.51	1.582	1.559	1.555	1.489
26	2001	1.472	1.484	1.447	1.564	1.729	1.64	1.482	1.427	1.531	1.362	1.263	1.131
27	2002	1.139	1.13	1.241	1.407	1.421	1.404	1.412	1.423	1.422	1.449	1.448	1.394
28	2003	1.473	1.641	1.748	1.659	1.542	1.514	1.524	1.628	1.728	1.603	1.535	1.494
29	2004	1.592	1.672	1.766	1.833	2.009	2.041	1.939	1.898	1.891	2.029	2.01	1.882
30	2005	1.823	1.918	2.065	2.283	2.216	2.176	2.316	2.506	2.927	2.785	2.343	2.186
31	2006	2.315	2.31	2.401	2.757	2.947	2.917	2.999	2.985	2.589	2.272	2.241	2.334
32	2007	2.274	2.285	2.592	2.86	3.13	3.052	2.961	2.782	2.789	2.793	3.069	3.02
33	2008	3.047	3.033	3.258	3.441	3.764	4.065	4.09	3.786	3.698	3.173	2.151	1.689
34	2009	1.787	1.928	1.949	2.056	2.265	2.631	2.543	2.627	2.574	2.561	2.66	2.621
35	2010	2.731	2.659	2.78	2.858	2.869	2.736	2.736	2.745	2.704	2.795	2.852	2.985
36	2011	3.091	3.167	3.546	3.816	3.933	3.702	3.654	3.63	3.612	3.468	3.423	3.278
37	2012	3.399	3.572	3.868	3.927	3.792	3.552	3.451	3.707	3.856	3.786	3.488	3.331
38	2013	3.351	3.693	3.735	3.59	3.623	3.633	3.628	3.6	3.556	3.375	3.251	3.277
39	2014	3.32	3.364	3.532	3.659	3.691	3.695	3.633	3.481	3.403	3.182	2.887	2.56
40	2015	2.11	2.249	2.483	2.485	2.775	2.832	2.832	2.679	2.394	2.289	2.185	2.06

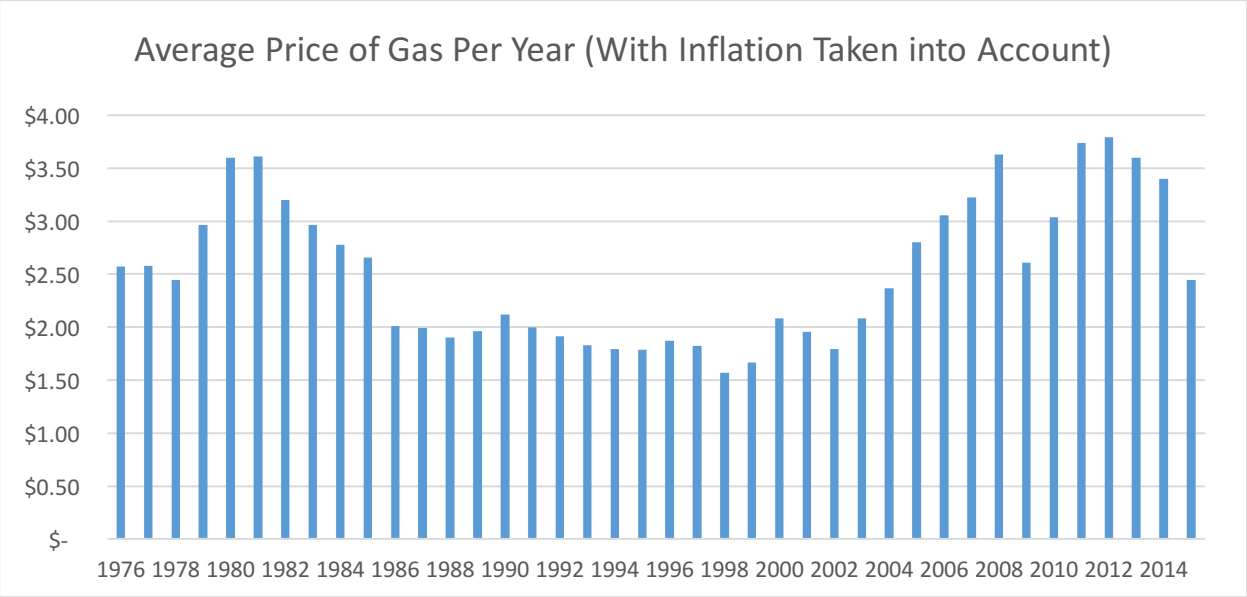
(Figure 2)

The entirety of this project (including source, presentation, and this report) can be found on GitHub: Here

Samuel Wallace
Connor McInnes
ECE102 Project Report

Cost Per Gallon of Gasoline														
Year	Month												Average Per Year	Average Per Year + Inflation
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec		
1976	$ 0.605	$ 0.600	$ 0.594	$ 0.592	$ 0.600	$ 0.616	$ 0.623	$ 0.628	$ 0.630	$ 0.629	$ 0.629	$ 0.626	$ 0.614	$ 2.57
1977	$ 0.627	$ 0.637	$ 0.643	$ 0.651	$ 0.659	$ 0.665	$ 0.667	$ 0.667	$ 0.666	$ 0.665	$ 0.664	$ 0.665	$ 0.656	$ 2.58
1978	$ 0.648	$ 0.647	$ 0.647	$ 0.649	$ 0.655	$ 0.663	$ 0.674	$ 0.682	$ 0.688	$ 0.690	$ 0.695	$ 0.705	$ 0.670	$ 2.45
1979	$ 0.716	$ 0.730	$ 0.755	$ 0.802	$ 0.844	$ 0.901	$ 0.949	$ 0.988	$ 1.020	$ 1.028	$ 1.041	$ 1.065	$ 0.903	$ 2.96
1980	$ 1.131	$ 1.207	$ 1.252	$ 1.264	$ 1.266	$ 1.269	$ 1.271	$ 1.267	$ 1.257	$ 1.250	$ 1.250	$ 1.258	$ 1.245	$ 3.60
1981	$ 1.298	$ 1.382	$ 1.417	$ 1.412	$ 1.400	$ 1.391	$ 1.382	$ 1.376	$ 1.376	$ 1.371	$ 1.369	$ 1.365	$ 1.378	$ 3.61
1982	$ 1.358	$ 1.334	$ 1.284	$ 1.225	$ 1.237	$ 1.309	$ 1.331	$ 1.323	$ 1.307	$ 1.295	$ 1.283	$ 1.260	$ 1.296	$ 3.20
1983	$ 1.230	$ 1.187	$ 1.152	$ 1.215	$ 1.259	$ 1.277	$ 1.288	$ 1.285	$ 1.274	$ 1.255	$ 1.241	$ 1.231	$ 1.241	$ 2.97
1984	$ 1.216	$ 1.209	$ 1.210	$ 1.227	$ 1.236	$ 1.229	$ 1.212	$ 1.196	$ 1.203	$ 1.209	$ 1.207	$ 1.193	$ 1.212	$ 2.78
1985	$ 1.148	$ 1.131	$ 1.159	$ 1.205	$ 1.231	$ 1.241	$ 1.242	$ 1.229	$ 1.216	$ 1.204	$ 1.207	$ 1.208	$ 1.202	$ 2.66
1986	$ 1.194	$ 1.120	$ 0.981	$ 0.888	$ 0.923	$ 0.955	$ 0.890	$ 0.843	$ 0.860	$ 0.831	$ 0.821	$ 0.823	$ 0.927	$ 2.01
1987	$ 0.862	$ 0.905	$ 0.912	$ 0.934	$ 0.941	$ 0.958	$ 0.971	$ 0.995	$ 0.990	$ 0.976	$ 0.976	$ 0.961	$ 0.948	$ 1.99
1988	$ 0.933	$ 0.913	$ 0.904	$ 0.930	$ 0.955	$ 0.955	$ 0.967	$ 0.987	$ 0.974	$ 0.957	$ 0.949	$ 0.930	$ 0.946	$ 1.90
1989	$ 0.918	$ 0.926	$ 0.940	$ 1.065	$ 1.119	$ 1.114	$ 1.092	$ 1.057	$ 1.029	$ 1.027	$ 0.999	$ 0.980	$ 1.022	$ 1.96
1990	$ 1.042	$ 1.037	$ 1.023	$ 1.044	$ 1.061	$ 1.088	$ 1.084	$ 1.190	$ 1.294	$ 1.378	$ 1.377	$ 1.354	$ 1.164	$ 2.12
1991	$ 1.247	$ 1.143	$ 1.082	$ 1.104	$ 1.156	$ 1.160	$ 1.127	$ 1.140	$ 1.143	$ 1.122	$ 1.134	$ 1.123	$ 1.140	$ 2.00
1992	$ 1.073	$ 1.054	$ 1.058	$ 1.079	$ 1.136	$ 1.179	$ 1.174	$ 1.158	$ 1.158	$ 1.154	$ 1.159	$ 1.136	$ 1.127	$ 1.92
1993	$ 1.117	$ 1.108	$ 1.098	$ 1.112	$ 1.129	$ 1.130	$ 1.109	$ 1.097	$ 1.085	$ 1.127	$ 1.113	$ 1.070	$ 1.108	$ 1.83
1994	$ 1.043	$ 1.051	$ 1.045	$ 1.064	$ 1.080	$ 1.106	$ 1.136	$ 1.182	$ 1.177	$ 1.152	$ 1.163	$ 1.143	$ 1.112	$ 1.79
1995	$ 1.129	$ 1.120	$ 1.115	$ 1.140	$ 1.200	$ 1.226	$ 1.195	$ 1.164	$ 1.148	$ 1.127	$ 1.101	$ 1.101	$ 1.147	$ 1.79
1996	$ 1.129	$ 1.124	$ 1.162	$ 1.251	$ 1.323	$ 1.299	$ 1.272	$ 1.240	$ 1.234	$ 1.227	$ 1.250	$ 1.260	$ 1.231	$ 1.87
1997	$ 1.261	$ 1.255	$ 1.235	$ 1.231	$ 1.226	$ 1.229	$ 1.205	$ 1.253	$ 1.277	$ 1.242	$ 1.213	$ 1.177	$ 1.234	$ 1.83
1998	$ 1.131	$ 1.082	$ 1.041	$ 1.052	$ 1.092	$ 1.094	$ 1.079	$ 1.052	$ 1.033	$ 1.042	$ 1.028	$ 0.986	$ 1.059	$ 1.57
1999	$ 0.972	$ 0.955	$ 0.991	$ 1.177	$ 1.178	$ 1.148	$ 1.189	$ 1.255	$ 1.280	$ 1.274	$ 1.264	$ 1.298	$ 1.165	$ 1.67
2000	$ 1.301	$ 1.369	$ 1.541	$ 1.506	$ 1.498	$ 1.617	$ 1.593	$ 1.510	$ 1.582	$ 1.559	$ 1.555	$ 1.489	$ 1.510	$ 2.08
2001	$ 1.472	$ 1.484	$ 1.447	$ 1.564	$ 1.729	$ 1.640	$ 1.482	$ 1.427	$ 1.531	$ 1.362	$ 1.263	$ 1.131	$ 1.461	$ 1.96
2002	$ 1.139	$ 1.130	$ 1.241	$ 1.407	$ 1.421	$ 1.404	$ 1.412	$ 1.423	$ 1.422	$ 1.449	$ 1.448	$ 1.394	$ 1.358	$ 1.79
2003	$ 1.473	$ 1.641	$ 1.748	$ 1.659	$ 1.542	$ 1.514	$ 1.524	$ 1.628	$ 1.728	$ 1.603	$ 1.535	$ 1.494	$ 1.591	$ 2.08
2004	$ 1.592	$ 1.672	$ 1.766	$ 1.833	$ 2.009	$ 2.041	$ 1.939	$ 1.898	$ 1.891	$ 2.029	$ 2.010	$ 1.882	$ 1.880	$ 2.37
2005	$ 1.823	$ 1.918	$ 2.065	$ 2.283	$ 2.216	$ 2.176	$ 2.316	$ 2.506	$ 2.927	$ 2.785	$ 2.343	$ 2.186	$ 2.295	$ 2.80
2006	$ 2.315	$ 2.310	$ 2.401	$ 2.757	$ 2.947	$ 2.917	$ 2.999	$ 2.985	$ 2.589	$ 2.272	$ 2.241	$ 2.334	$ 2.589	$ 3.05
2007	$ 2.274	$ 2.285	$ 2.592	$ 2.860	$ 3.130	$ 3.052	$ 2.961	$ 2.782	$ 2.789	$ 2.793	$ 3.069	$ 3.020	$ 2.801	$ 3.22
2008	$ 3.047	$ 3.033	$ 3.258	$ 3.441	$ 3.764	$ 4.065	$ 4.090	$ 3.786	$ 3.698	$ 3.173	$ 2.151	$ 1.689	$ 3.266	$ 3.63
2009	$ 1.787	$ 1.928	$ 1.949	$ 2.056	$ 2.265	$ 2.631	$ 2.543	$ 2.627	$ 2.574	$ 2.561	$ 2.660	$ 2.621	$ 2.350	$ 2.61
2010	$ 2.731	$ 2.659	$ 2.780	$ 2.858	$ 2.869	$ 2.736	$ 2.736	$ 2.745	$ 2.704	$ 2.795	$ 2.852	$ 2.985	$ 2.788	$ 3.04
2011	$ 3.091	$ 3.167	$ 3.546	$ 3.816	$ 3.933	$ 3.702	$ 3.654	$ 3.630	$ 3.612	$ 3.468	$ 3.423	$ 3.278	$ 3.527	$ 3.74
2012	$ 3.399	$ 3.572	$ 3.868	$ 3.927	$ 3.792	$ 3.552	$ 3.451	$ 3.707	$ 3.856	$ 3.786	$ 3.488	$ 3.331	$ 3.644	$ 3.79
2013	$ 3.351	$ 3.693	$ 3.735	$ 3.590	$ 3.623	$ 3.633	$ 3.628	$ 3.600	$ 3.556	$ 3.375	$ 3.251	$ 3.277	$ 3.526	$ 3.60
2014	$ 3.320	$ 3.364	$ 3.532	$ 3.659	$ 3.691	$ 3.695	$ 3.633	$ 3.481	$ 3.403	$ 3.182	$ 2.887	$ 2.560	$ 3.367	$ 3.40
2015	$ 2.110	$ 2.249	$ 2.483	$ 2.485	$ 2.775	$ 2.832	$ 2.832	$ 2.679	$ 2.394	$ 2.289	$ 2.185	$ 2.060	$ 2.448	$ 2.45
Cheapest Year To Buy Gas	1998													

(Figure 3)



(Figure 4)

The entirety of this project (including source, presentation, and this report) can be found on GitHub: Here

As can be seen in figure 4, once we have imported the data into an excel spreadsheet, it is also quite trivial to manipulate the said data into a charts which convey entirely new information than what was original to the source. We are quite pleased with the results of this project, and flexibility that is offered through the reformatting of our original data. As we have open sourced our code, our base work here could be expanded upon by others to allow refactoring of other databases into a more flexible formats, as the logic of moving from html source to xml to excel remains the same. We hope it might prove useful to someone in the future.

The entirety of this project (including source, presentation, and this report) can be found on GitHub: Here