SESUG Paper 249-2018

Worker Safety in Energy Production in America
A Comparative Analysis

Alyssa Venn, Kennesaw State University

Joe DeMaio, Kennesaw State University

ABSTRACT

This research looks at injuries, fatal and non-fatal, and energy production in several American energy fields, comparing different resources on standardized scales as well as how injury rates have changed through the years. Discussions of the safety of coal vs. other energy sources have been heated, especially recently, and this research project analyzes injury rates in an analytical, objective manner.

INTRODUCTION

The world runs on energy. From light bulbs to cell phones to cars, energy is essential. In our increasingly technology-based society, its importance, too, is increasing. As our reliance on energy increases, the need to find safe and sustainable energy production sources only becomes more urgent. In the United States, as of 2016, coal energy makes up about 30% of electricity generated, while hydroelectric provides about 7% (U.S. Energy Information Administration, 2017). Coal has a reputation as one of the more dangerous fields to work in. In 2007, the rate of fatal injuries for coal miners was almost six times the rate of fatal injuries in private industry (Bureau of Labor Statistics, 2010).

METHODS

This project gathers data on injuries (fatal and non-fatal), as well as total employees employed in each industry, from the U.S. Bureau of Labor Statistics' website, and data on energy produced from the U.S. Energy Information Administration. The data looks at coal, hydroelectric, petroleum, and drilling from 1992 to 2016. "Drilling" was made up of two fields–natural gas extraction and oil extraction–as the sources for energy combined the two. The data came in both PDF and XLSX format, which was cleaned in Excel and SAS. The data were analyzed with SAS 9.04.

WORKER SAFETY IN COAL

Figure 1 depicts the time series of non-fatal injuries resulting in 31 or more days of work missed (the blue scatter plot curve) in the coal industry, as well as a prediction model (the red curve). It can be seen that the model is logarithmic, fitting the equation $f(t) = e^{122.08 - 0.0572t}$, which decreases as time goes on. This indicates that the number of injuries in the coal industry is being reduced. This model predicts that in 2030, there will be approximately 390 nonfatal injuries resulting in workers missing 31 or more days of work. The Box-Cox transform indicated that the logarithmic model was the best fit. Code is available in Appendix I.

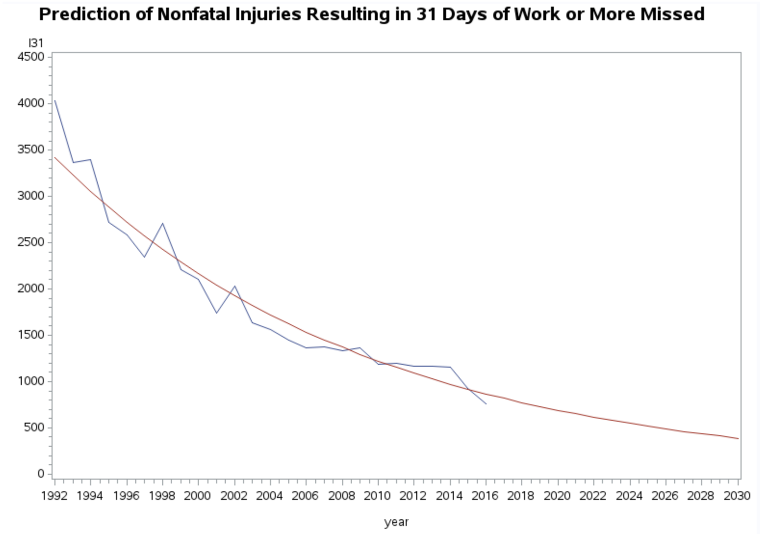

Figure 1. Prediction of Nonfatal Injuries in the Coal Industry Resulting in 31 Days of Work or More Missed

However, just looking at the raw numbers of nonfatal injuries wouldn't tell the whole story, as it's possible that the coal industry is shrinking which results in decreasing the number of injuries. Figure 2 shows the raw numbers of thousand workers in the coal industry by year, and Figure 3 standardizes fatalities per 1000 workers. Figure 3 shows the ratio of injuries per thousand workers, with a smaller number meaning fewer people are being injured relative to the number of workers. As seen in the time series, the ratio is also decreasing as the years go by, meaning that there are fewer people being severely injured relative to the number of people working in the field.

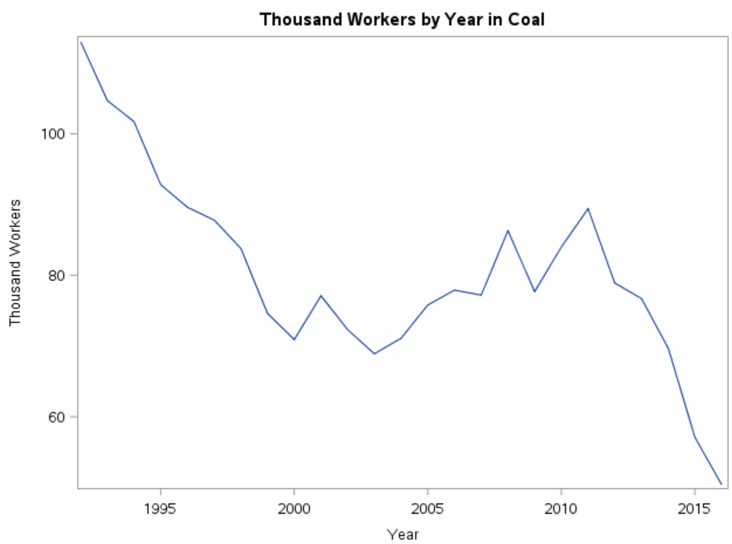

Figure 2. Thousand Workers in the Coal Industry by Year

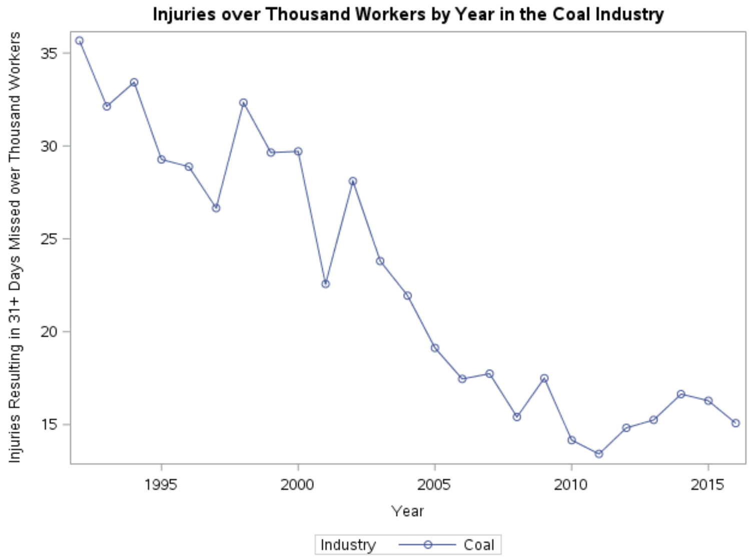

Figure 3. Injuries Resulting in 31+ Days of Work Missed over Thousand Workers in the Coal Industry by Year

Figure 4 shows a similar trend, only fatalities instead of injuries. With fatalities as well, though the graph is more erratic than Figure 3, fatalities in the coal industry seem to be decreasing even when taking the decreasing workforce into account. Figure 5 relates injuries to fatalities, showing how many people are severely injured per fatality in coal mining. The graph is rather inconsistent, showing no real relation between serious injuries and fatalities. Perhaps safety efforts have been more strongly focused on preventing fatalities rather than serious injury.

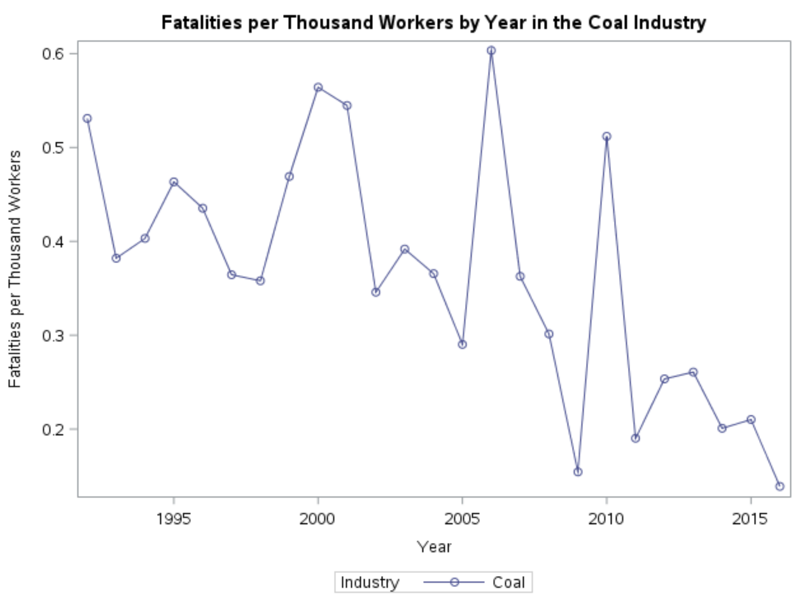

Figure 4. Fatalities per Thousand Workers in the Coal Industry by Year

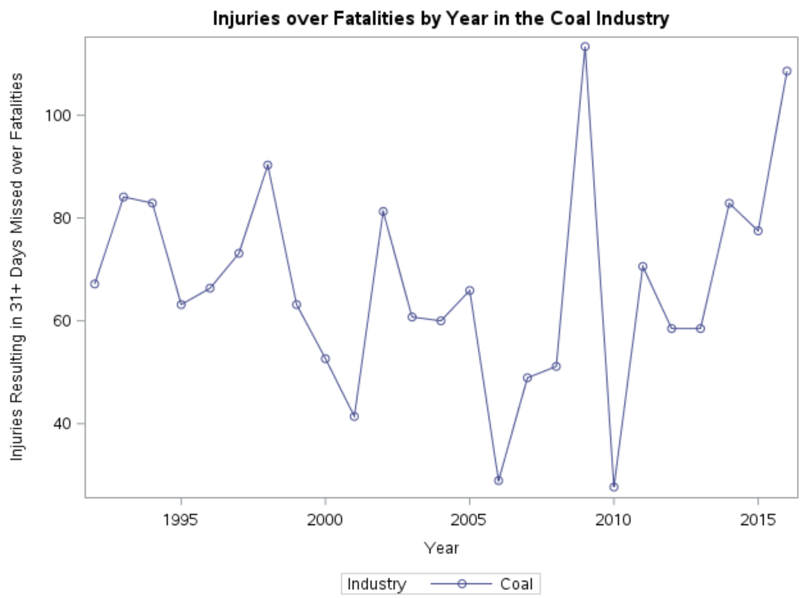

Figure 5. Injuries Resulting in 31+ Days of Work Missed per Fatality in the Coal Industry by Year

COMPARING WORKER SAFETY BETWEEN FIELDS

Figure 6 depicts raw fatalities by industry through the years. Here we can see that the raw number of fatalities for coal and drilling for oil and natural gas are similar, with petroleum being, on average, lower. No one field is always safer or more hazardous than the others, save hydroelectric, when looking at the raw numbers.

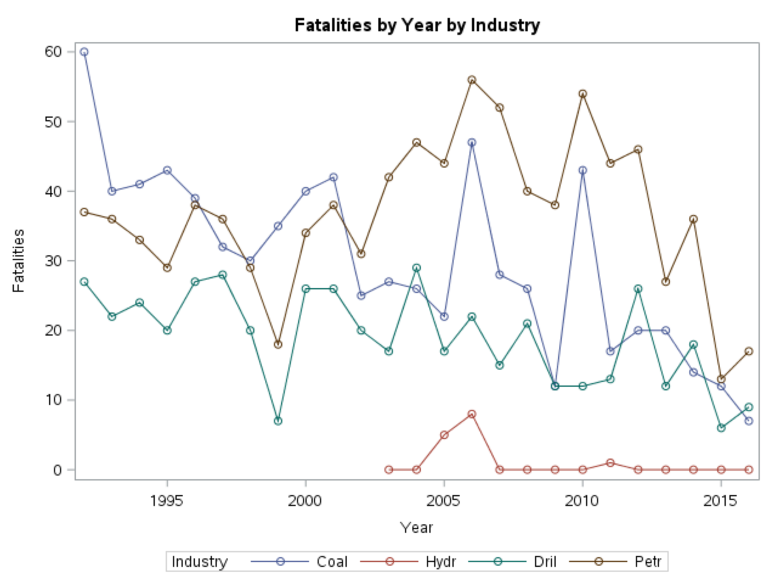

Figure 6. Fatalities by Year by Industry

However, this again may not tell the whole story. We standardized these numbers using two methods: per thousand workers employed in the field and by energy produced. In standardizing by energy produced, we are looking at how many deaths it takes to produce one million kilowatt-hours. Standardizing by workers tells us approximately how dangerous it is to work in that field, as any singular worker.

Figure 7 shows energy produced, standardized by the number of workers in each field. Coal, petroleum, and drilling are all comparably stable, with coal seeming to be more efficient with regards to workers than both drilling and petroleum. Of note is that though hydroelectric power begins less efficient than coal, it does not stay that way. There is a clear spike in energy produced relative to workers in 2009, peaking in 2010 but still very high afterwards. This may be attributed to the American Recovery and Reinvestment Act of 2009, which sponsored the Water Power Program, intending to increase efficiency of hydropower. To do so, the program sponsored a team of U.S. scientists to develop tools to help managers and operators to more-efficiently run their hydropower plants (Office of Energy Efficiency & Renewable Energy, n.d.). If this graph is anything to go by, they were successful, as it appears that energy produced changed drastically, without needing to increase the number of workers. On the other hand, petroleum is almost completely flat, implying that they are not becoming any more efficient, while drilling appears to be becoming very gradually more efficient.

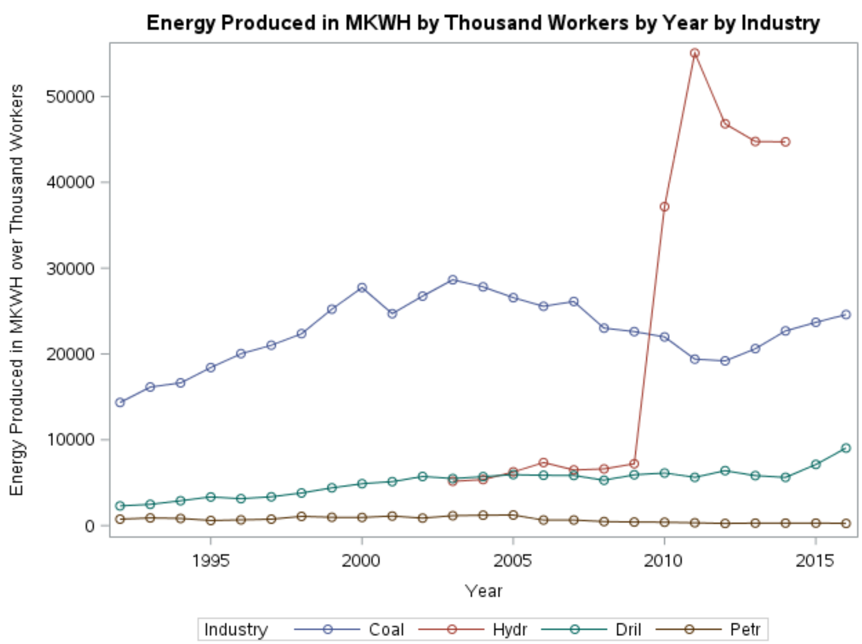

Figure 7. Energy Produced in Million Kilowatt-Hours by Thousand Workers by Year by Industry

STANDARDIZING BY ENERGY PRODUCED

Figure 8 depicts the fatalities in coal, hydroelectric, drilling, and crude petroleum extraction, standardized by energy produced in that field. Because the fields are so different with regards to size, the data were analyzed as a ratio of how many people died for each million kilowatt hours of energy produced. It can be seen that petroleum is drastically higher than the other fields, as well as being very erratic. There is a large spike in relative fatalities between 2006 and 2011, and though petroleum fatalities seem to be trending down more recently, they are still higher than they were previous to 2005. This is very different from the raw numbers, in which petroleum is generally the least dangerous of the top three. This implies that petroleum does not create very much energy, especially considering frequency of fatalities.

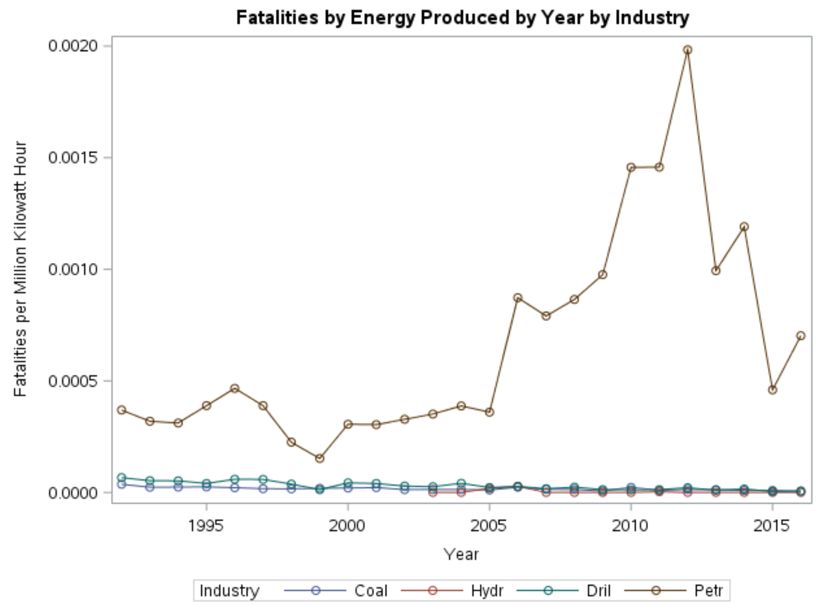

Figure 8. Fatalities by Energy Produced by Year by Industry

Figure 9 removes petroleum so that we can drill down to examine the other fields more closely. Coal and drilling relative fatalities do seem to be gradually decreasing, while hydroelectric is unpredictable. Because hydroelectric generally has no fatalities, it is impossible to model how it is changing. However, though coal is generally assumed to be much more dangerous, there are two years–2005 and 2006– where the relative fatalities in hydroelectric energy generation exceeds those in coal. This implies that, though hydroelectric is safer on average, when something goes wrong, it is catastrophic.

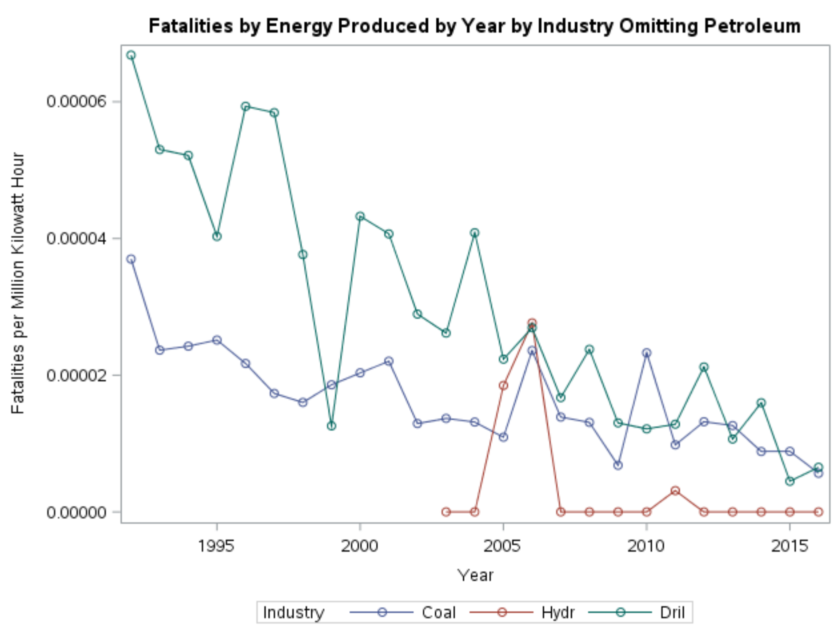

Figure 9. Fatalities by Energy Produced by Year by Industry Omitting Petroleum

Analyzing graphs of data trends can provide insight into the occupational risk in these industries.

However, are these differences significant? A one-way between subjects ANOVA was conducted to compare the effect of industry on fatalities by energy produced for coal, hydroelectric, drilling, and petroleum. There was a significant effect of industry on fatalities by energy produced at the $p<.05$ level for the four conditions [$F(3, 85) = 39.48$, $p < 0.0001$]. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the petroleum condition (M = 6.56E-4, SD = 4.67E-4) was significantly different than the coal condition (M = 1.69E-5, SD = 7.25E-6), the drilling condition (M = 2.99E-5, SD = 1.82E-5), and the hydroelectric condition (M = 3.52E-6, SD = 8.52E-6). However, the coal, drilling, and hydroelectric conditions did not differ from each other. Figure 10 shows side-by-side boxplots of of the industries, showing the differences.

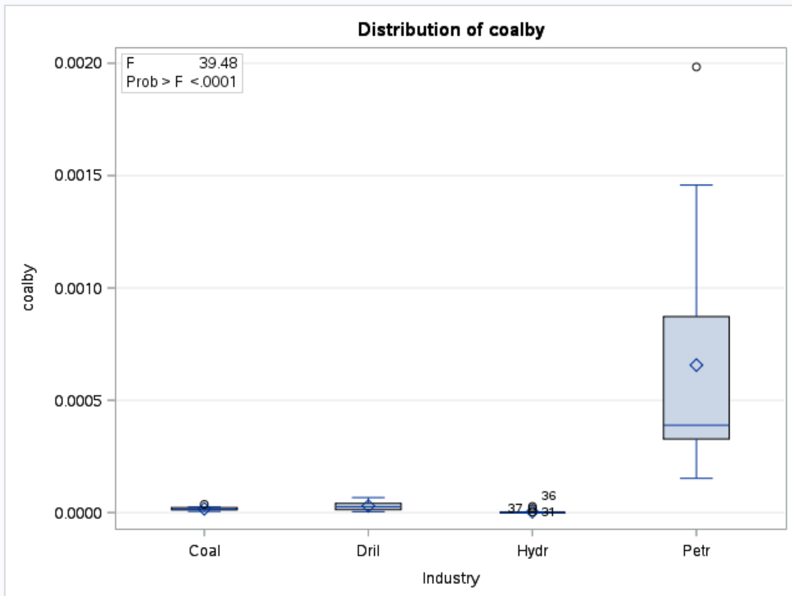

Figure 10. Side-by-Side Boxplots of Ratio of Energy Produced to Fatalities by Industry

Taken together, these results suggest that the petroleum field is significantly more dangerous than the coal, hydroelectric, and drilling extraction fields, when standardizing by energy produced. The other three energy fields show no significant difference between each other. The results of the ANOVA and consequent Tukey HSD test can be found in Appendix II.

STANDARDIZING BY WORKERS

Though hydroelectric, coal, and drilling for gas and oil have approximately the same pattern when standardized by workers as when standardized by energy, as seen in Figure 10, petroleum is very different. While petroleum by energy gets more erratic as the years go on, petroleum by workers doesn't have quite as many spikes and dips, instead forming a sort of bell shape. Petroleum by workers actually appears to be improving. There is also a difference in hydroelectric and coal, as hydroelectric fatalities standardized by energy produced exceed those of coal in two years, but hydroelectric fatalities standardized by number of workers is always less than coal's.

Figure 11. Fatalities by Thousand Workers by Year by Industry

A one-way between subjects ANOVA was conducted to compare the effect of industry on fatalities by 1000 workers employed for coal, hydroelectric, drilling, and petroleum. There was a significant effect of industry on fatalities 1000 workers at the p<.05 level for the four conditions [F(3, 83) = 43.57, p < 0.0001]. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the petroleum condition (M = 0.3457, SD = 0.1231) and the coal condition (M = 0.3639, SD = 0.1317) were significantly different than the drilling condition (M = 0.1292, SD = 0.0544) and the hydroelectric condition (M = 0.0409, SD = 0.0764). However, the coal and petroleum conditions were not significantly different from each other, and the drilling and hydroelectric conditions also did not differ from each other. Figure 12 shows side-by-side boxplots of the industries, demonstrating the differences.

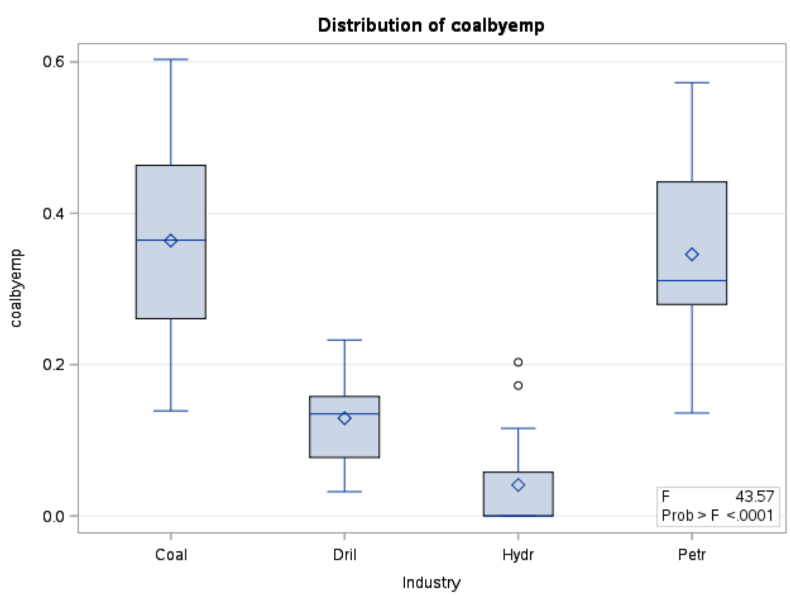

Figure 12. Side-by-Side Boxplots of Ratio of Fatalities to Total Employees by Industry

Taken together, these results suggest that the petroleum and coal fields are significantly more dangerous than the hydroelectric and drilling fields, when standardizing by 1000 employees. Coal and petroleum, however, have no significant difference, nor do hydroelectric and drilling. The results of the ANOVA and consequent Tukey HSD test can be found in Appendix III.

CONCLUSION

In summary, we found that the data is not as simple as "coal is bad." We were surprised to see that, despite coal having what is seemingly the worst reputation, it is not the worst with regards to fatalities. While both fatalities and nonfatal injuries are steadily decreasing in the coal industry, this is not the same with the others. Hydroelectric, drilling, and petroleum are too variable to see any real pattern. When standardized by energy produced, petroleum is drastically more dangerous than the other three, with coal, drilling, and hydroelectric not differing significantly from each other. With regards to fatalities by total workers, coal and petroleum trade off for most dangerous, with no significant difference between the two of them, though they are both significantly different from hydroelectric and drilling.

REFERENCES

Bureau of Labor Statistics (2010, April). Injuries, illnesses, and fatalities in the coal mining industry. Retrieved from https://www.bls.gov/iif/oshwc/osh/os/osar0012.htm.

Kuhfeld, W. F. (2015). Advanced ODS graphics examples. Cary, NC: SAS Institute Inc.

Kuhfeld, W. F. (2016). Basic ODS graphics examples. Cary, NC: SAS Institute Inc.

Office of Energy Efficiency & Renewable Energy. (n.d.). Hydropower technology development, *Energy.gov*. Retrieved from https://www.energy.gov/eere/water/hydropower-technology-development

U.S. Energy Information Administration (2017, April 4). Total energy. Retrieved from https://www.eia.gov/totalenergy/data/annual/index.php

U.S. Energy Information Administration (2017, May 10). Energy explained: Electricity in the United States. Retrieved from https://www.eia.gov/energyexplained/index.cfm?page=electricity_in_the_united_states

ACKNOWLEDGMENTS

This work is not an island entirely of itself. Many groups supported this work through its various stages. A big thank you goes to Birla Carbon for providing funding to support this work. Thanks go to the College of Science and Mathematics at Kennesaw State University for its relationship and hosting of the Birla Carbon Scholars Program. Thank you to the folks at the SAS Support Communities for their help with coding tips. Thanks go to the Analytics and Data Science Institute at Kennesaw State University for providing computing resources and allowing us to present preliminary results at their annual Analytics Day. Thanks also go to Kennesaw State University for hosting the Undergraduate Research Symposium to showcase this work. Finally, another big thank you goes to the Southeast SAS Users Group (SESUG) for publishing this work in their conference proceedings as well as providing travel support in the form of an Academic Student grant.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Alyssa Venn
avenn1@students.kennesaw.edu

Joe DeMaio
jdemaio@kennesaw.edu

APPENDIX I

```
proc transreg details data=coal2;
model boxcox("31+ number"n) = identity(year);
run;

data new;
set coal2;
ln31 = log("31+ number"n);
run;
proc reg data=new;
model ln31=year / cli;
run;
data function;
do year = 1992 to 2030 by 1;
   amnt = exp(122.08-0.0572*year);
   output;
   end;
run;
symbol1 i=join;
footnote1 h=1 c=green;
footnote3 h=1 c=blue;
proc gplot data=coal4;
title "Prediction of Nonfatal Injuries Resulting in 31 Days of Work or More
Missed";
plot l31*year amnt*year/ overlay
haxis=1992 to 2030 by 2 vaxis=0 to 4500 by 500;
run;
                                  quit;
```

APPENDIX II

The ANOVA Procedure

Dependent Variable: coalby

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.00000731	0.00000244	39.48	<.0001
Error	85	0.00000525	0.00000006		
Corrected Total	88	0.00001256			

R-Square	Coeff Var	Root MSE	coalby Mean
0.582198	125.4764	0.000248	0.000198

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Industry	3	7.311996E-6	2.437332E-6	39.48	<.0001

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for coalby

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	85
Error Mean Square	6.173E-8
Critical Value of Studentized Range	3.70608

Comparisons significant at the 0.05 level are indicated by ***.

Industry Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
Petr - Dril	0.00062655	0.00044238	0.00081071	***
Petr - Coal	0.00063976	0.00045559	0.00082392	***
Petr - Hydr	0.00065290	0.00043555	0.00087025	***
Dril - Petr	-.00062655	-.00081071	-.00044238	***
Dril - Coal	0.00001321	-.00017095	0.00019737	
Dril - Hydr	0.00002635	-.00019099	0.00024370	
Coal - Petr	-.00063976	-.00082392	-.00045559	***
Coal - Dril	-.00001321	-.00019737	0.00017095	
Coal - Hydr	0.00001314	-.00020420	0.00023049	
Hydr - Petr	-.00065290	-.00087025	-.00043555	***
Hydr - Dril	-.00002635	-.00024370	0.00019099	
Hydr - Coal	-.00001314	-.00023049	0.00020420	

APPENDIX III

The ANOVA Procedure

Dependent Variable: coalbyemp

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1.44145971	0.48048657	43.57	<.0001
Error	83	0.91525689	0.01102719		
Corrected Total	86	2.35671660			

R-Square	Coeff Var	Root MSE	coalbyemp Mean
0.611639	42.56739	0.105010	0.246692

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Industry	3	1.44145971	0.48048657	43.57	<.0001

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for coalbyemp

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	83
Error Mean Square	0.011027
Critical Value of Studentized Range	3.70786

Comparisons significant at the 0.05 level are indicated by ***.

Industry Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
Coal - Petr	0.01814	-0.05973	0.09602	
Coal - Dril	0.23464	0.15677	0.31251	***
Coal - Hydr	0.32294	0.22625	0.41963	***
Petr - Coal	-0.01814	-0.09602	0.05973	
Petr - Dril	0.21650	0.13862	0.29437	***
Petr - Hydr	0.30480	0.20811	0.40149	***
Dril - Coal	-0.23464	-0.31251	-0.15677	***
Dril - Petr	-0.21650	-0.29437	-0.13862	***
Dril - Hydr	0.08830	-0.00839	0.18499	
Hydr - Coal	-0.32294	-0.41963	-0.22625	***
Hydr - Petr	-0.30480	-0.40149	-0.20811	***
Hydr - Dril	-0.08830	-0.18499	0.00839	