Energy Generation Consumption Review

Christopher Pamieta
christopherpamieta@lewisu.edu

Jason Yakubovich
jasonayakubovich@lewisu.edu

Dumisani Dlamini
dumisanidlamini@lewisu.edu

Michael Sloan
MichaelDSloan@lewisu.edu

I. Introduction

This report (The Energy Consumption Review) serves two purposes. First, it describes a series of datasets provided by the U.S. Energy Information Administration and speaks convincingly towards developing trends in energy consumption and production for the United States. Second, it serves as a learning vehicle by which students of the CPSC53000 – Data Visualization Course demonstrate the selection of visualization techniques and analyze the resulting techniques used to build effective visualizations.

We chose an energy dataset, and what motivated the group to choose this dataset was the keen interest in understanding some of the underlying drivers for power consumption. In other words, the team wanted to find out if there was any direct relationship between energy use and production on the different energy types.

II. Datasets

The U.S Energy Information Administration (USEIA) is the United States governmental agency organized to collect, analyze, and disseminate energy information on behalf of the United States government. The USEIA produces a series of datasets as part of its overarching Monthly Energy Review [1]. These datasets contain recent and historical time-series data measuring energy consumption and production categorized by various attributes including consumption sectors and production types.

The dataset contains energy consumption by residential, commercial, industrial, transportation, and power sectors. The well-organized data format required minimal cleaning. Several rows of column headings required deletion, and non-data rows between the headings and the data required removal. Each sector contains primary energy consumed and total energy consumed data. The energy dataset spans a period from January 1973 up to July 2016 and uses units of trillion BTU (British thermal units). The datatype for the energy dataset is categorical data since it is "organized" or categorized by sector, which, in other terms, are different consumption categories that will be explored and analyzed. Due to the structure of the complimentary dataset, the Tableau shaper plugin for Excel was used to move all of the columns to rows. This resulted in a dataset with five columns, while the original dataset had thirteen columns.

The complimentary data set contains energy production by type [2]. The energy production types include coal, oil, natural gas, the wind, hydroelectric, and nuclear. This data used quadrillion BTUs as a unit of measurement for power consumed. However, preprocessing converted it to trillion BTUs by multiplying by a thousand. This data listed all of the categories under a single type column with the corresponding measures in a value column instead of each category having its column.

The original complimentary data set was transformed this to mimic prior set and was simply copied and pasted to the bottom to form the complete data set. Calculated Fields in Tableau were then used to get the specific information for each visualization.

III. Raw Dataset Schemas

Table 1 represents the original, untransformed dataset for Data Table 2.1 – Energy Consumption by Sector. The order of the schema matches the order of the columns in the raw dataset.

Table 2 represents the original, untransformed dataset for Data Table 1.2 – Primary Energy Production by Source. The order of the schema matches the order of the columns in the raw dataset.

Table 3 shows the classification of the sources in Data Table 1.3 – Primary Energy Production by Source.

Table 1 – Dataset Table 2.1 – Sector Energy Consumption.

Field Name	Schema Metadata		
	Type	*Category*	*Purpose*
Month	Date - DDMMYYYY	Dimension	First Day of Data and Month for which the observation applies, Unique
Energy Consumption Balancing Item (Trillion BTU)	Numeric – Continuous	Measure	Primary energy consumption less the total energy consumption for the observation period.
Primary Energy Consumed by the Commercial Sector (Trillion BTU)	Numeric – Continuous	Measure	Primary energy consumed by the Commercial Sector for the observation period. Primary energy is the energy consumed in Trillions of BTUs to generate the energy consumed at the point of generation [3].
Primary Energy Consumed by the Electric Power Sector (Trillion BTU)	Numeric – Continuous	Measure	Primary energy consumed by the Electric Power Sector for the observation period.
Primary Energy Consumed by the Industrial Sector (Trillion BTU)	Numeric – Continuous	Measure	Primary energy consumed by the Industrial Sector for the observation period.
Primary Energy Consumed by the Residential Sector (Trillion BTU)	Numeric – Continuous	Measure	Primary energy consumed by the Residential Sector for the observation period.
Primary Energy Consumed by the Transportation Sector (Trillion BTU)	Numeric – Continuous	Measure	Primary energy consumed by the Transportation Sector for the observation period.
Primary Energy Consumption Total (Trillion BTU)	Numeric – Continuous	Measure	Total primary energy consumed across all sectors for the observation period.
Total Energy Consumed by the Commercial Sector (Trillion BTU)	Numeric – Continuous	Measure	Total energy consumed by the Commercial Sector for the observation period. Total energy is the energy consumed in Trillions of BTUs at the point of consumption [3].
Total Energy Consumed by the Electric Power Sector (Trillion BTU)	Numeric – Continuous	Measure	Total energy consumed by the Electric Power Sector for the observation period.

Total Energy Consumed by the Industrial Sector (Trillion BTU)	Numeric – Continuous	Measure	Total energy consumed by the Industrial Sector for the observation period.
Total Energy Consumed by the Residential Sector (Trillion BTU)	Numeric – Continuous	Measure	Total energy consumed by the Residential Sector for the observation period.
Total Energy Consumed by the Transportation Sector (Trillion BTU)	Numeric – Continuous	Measure	Total energy consumed by the Transportation Sector for the observation period.
Total Energy Consumption Total (Trillion BTU)	Numeric – Continuous	Measure	The sum of all Total energy consumed across all sectors for the observation period.

Table 2 – Dataset Table 1.2 – Primary Energy Production

Field Name	Schema Metadata		
	Type	*Category*	*Purpose*
MSN	String – Classifier	Dimension	The code indicates the value classification. Not used.
YYYYMM	String – Classifier	Dimension	Year and Month of the observation stored as a string YYYYMM.
Value	Numeric – Continuous	Measure	Primary energy (point of generation) produced for the observation.
Description	String – Classifier	Dimension	Description of the generation method, primary classifier – see Table 3 for additional detail.
Unit	String – Classifier	Dimension	Unit of measure for the Value.

Table 3 – Dataset Table 1.2 – Primary Energy Production – Description Classifiers

Classifier	Classifier Description	
	Component of	*Category*
Biomass Energy Production	Renewable	Energy produced from biomass (i.e., biofuels).
Coal Production	Fossil Fuel	Energy produced from coal sources.
Crude Oil Production	Fossil Fuel	Energy produced because of crude oil production.
Geothermic Energy Production	Renewable	Energy produced from geothermal sources.
Hydroelectric Power Production	Renewable	Energy produced by hydroelectric power plants.
Natural Gas (Dry) Production	Fossil Fuel	Energy produced from Natural gas in dry form.
Natural Gas Plan Liquids Production	Fossil Fuel	Energy produced form liquid natural gas.

Nuclear Electric Power Production	Nuclear Electric Power Production	Energy produced by nuclear power plants using fissionable materials.
Total Fossil Fuels Production	Summary	The sum of all values that are a component of fossil fuels.
Total Primary Energy Production	Summary	Sum of Total Fossil Fuels Production, Total Renewable Energy Productions, and Nuclear Electric Power Production
Total Renewable Energy Production	Summary	The sum of all values that are a component of renewable sources.
Wind Energy Production	Renewable	Energy produced by from geothermal resources.

IV. VISUALIZATION AND DESCRIPTIONS

A. Background for Figures and Descriptions

This report presents a series of data visualizations based on the raw datasets described above and subject to the transformations previously described. This report uses the following convention for presenting these visualizations:

- The report assigns each visualization to its section

- Each section starts on a new page.

- The report visualization displays the figure at the start of the section

- The first section of text describes the findings illustrated by the visualization.

- The next section describes the purpose behind the selection of the visualization and the encodings considered when producing the visualization.

B. Correspondence between the Tableau Report [4] and The Energy Consumption Review

This data used in this report consists of time-series data. Because the nature of the dataset is driven by time, the selection of graphics supports time-series analysis and explanation.

The Tableau Report contains visualizations not present in this document. The Tableau Report contains all generated visuals and exists as a superset of the visualizations in this document. This is intentional.

The Energy Consumption Review contains a selection of static visualizations developed from the exploratory visualizations created by the team. Based on these observations, the team constructed this report.

C. Visualizations Selected

Table 4 – Visualizations Selected for the Static Report

Visualization	Description	
	Type	*Purpose*
Yearly Total Energy Consumption	Line	Demonstrates the growth of total energy consumption with some marked declines.
Energy Consumption 1997-2015	Line	This visualization demonstrates the leveling and decline in energy consumption with annotated economic events.
Energy Consumption by Sector as a Percent of Total	Area	Demonstrates the proportion of energy consumption by sectors.
Yearly Total Energy Consumption by Month	Line	Shows the annual seasonal impact on consumption for the two categories Residential and Commercial Usage.
Consumed vs. Produced Energy	Line	Compares the growth of consumed and produced energy, demonstrating core growth trends.
Energy Production by Type	Line	Shows a broad comparison of Nuclear, Renewable, and Fossil Fuel production over time.
Percent of Fossil Fuels Production by Type	Line	Demonstrates the chances in the Fossil Fuel mix over time.

Average Monthly Production by Type (Trillion BTUs)	Circle	A quick comparison of the total power mix averaged over times.

D. Visualizations (Starts on Next Page)
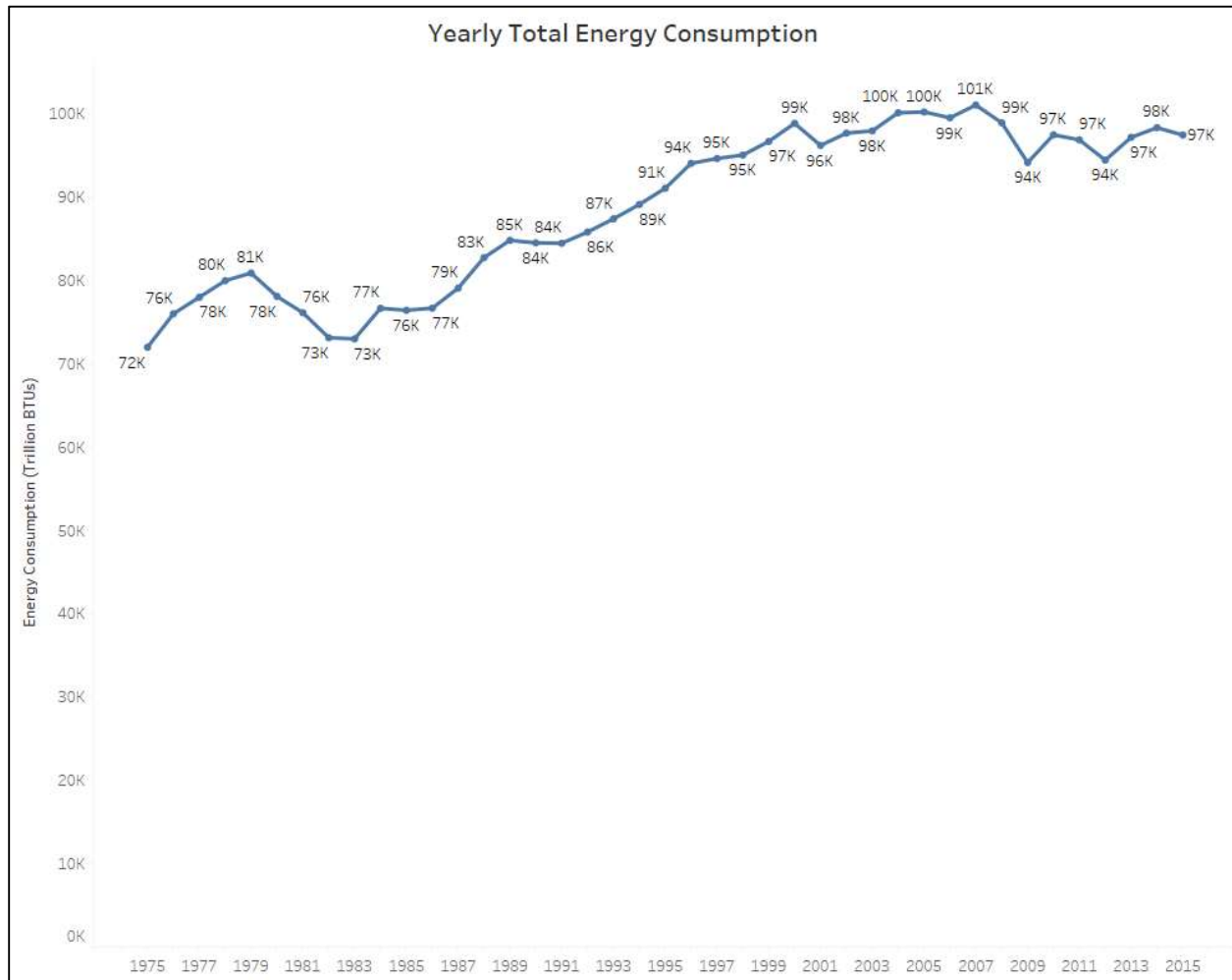
1) Total Energy Consumption

Figure 1 - Yearly Total Energy Consumption

a) Findings

Fig. 1 displays a generalized growth for annual energy consumption. However, marked declines start from highs in 1979 and 2001. The reader can interpret this graph to show that power consumption grew consistently from 1983 through 2008-2009, where consumption slightly declined and leveled off.

b) Intent of Graph

This visualization sets the starting point for the reader and lays the base assumption for the report. Consumption of power and growth remains largely consistent with a couple of marked downturns in that growth. Substantial declines in production started in 1979 and 2007. The graph asks the reader to answer "why."

c) Design Selections –

- The line graph was selected because it provides a mechanism for showing continuous data over a time-series.

- The blue color is used as the default for the line and the white background provide contrast and are not likely to confound a color-blind reader.

- Unit information is provided as part of the Y-Axis label (proximity), so it is clear as to its association with the Y-value.

- The labels on the line provide clarity of the values without cluttering the screen with gridlines or prescribing the use of a different style of graph.
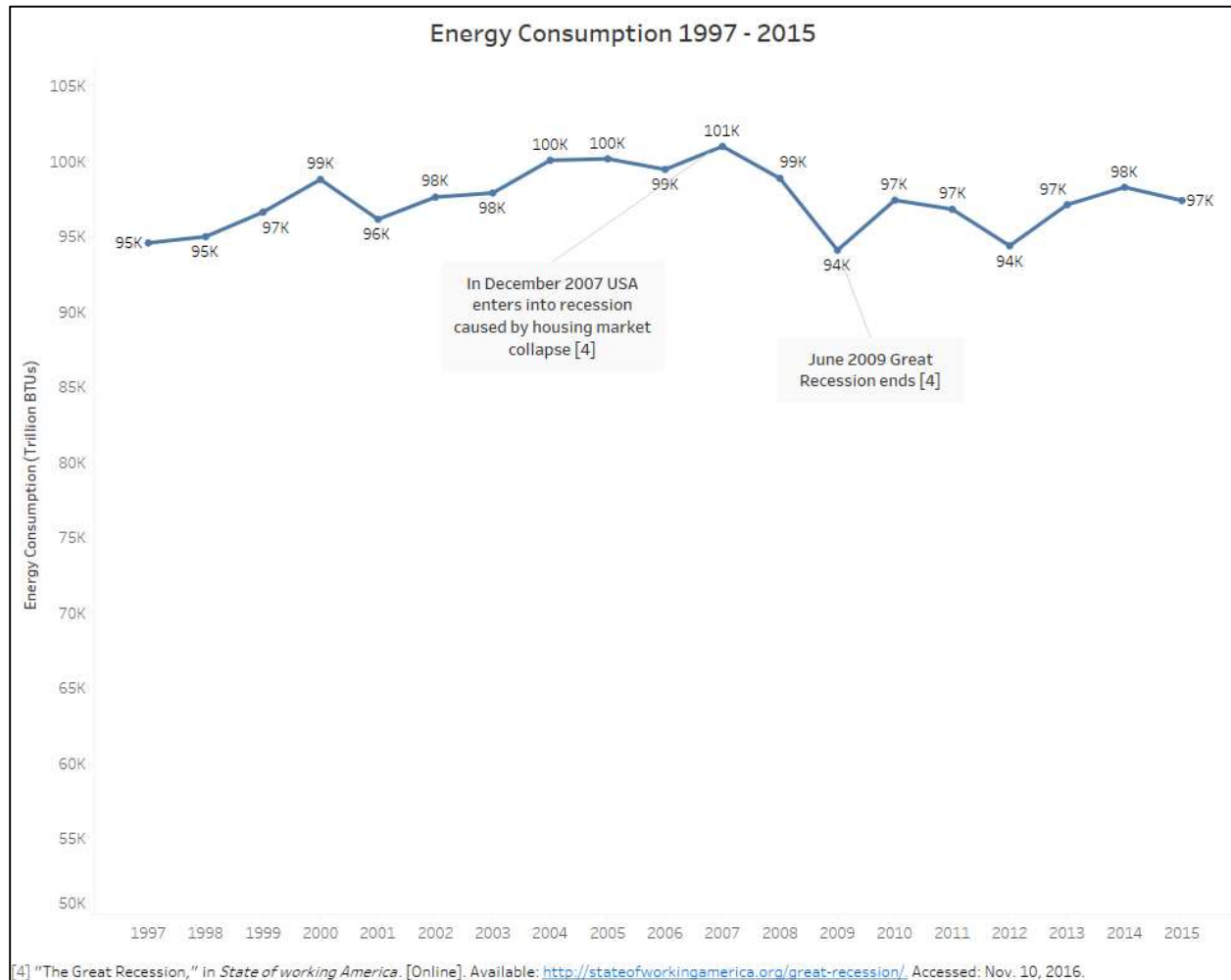
2) Energy Consumption 1997-2015

Figure 2 - Energy Consumption 1997-2015

   a) Findings

Fig. 2 reduces the view of the generalized growth for annual energy consumption shown in the previous visualization in Fig. 1. Overlaid events indicate the timing of the "Great Recession" for the United States.

   b) Intent of Graph

This visualization draws the user to the time series displayed before in the previous visualization and extended the narrative that something significant triggered these declines. The design selections lend continuity to the story, "Something happened" with "This could be what happened," later visualization that reinforces the message.

   c) Design Selections –

As this visualization extends from the original visualization in Fig 1., the design decision remains the same, with the addition of annotation.

- Annotation of events provides a visual description of the general causes of reduced power consumption. Because this balances against ample white space, it draws the eye to the message. This element serves as a replacement for aural information as an annotation.

- Lines from the annotation to the timed event value (not the year) associate and connect the data point with the event.

- The following design selections carry over from the previous figure (Fig. 1).

  o The line graph was selected because it provides a mechanism for showing continuous data over a time-series.

  o The blue color is used as the default for the line and the white background provide contrast and are not likely to confound a color-blind reader.

  o Unit information is provided as part of the Y-Axis label (proximity), so it is clear as to its association with the Y-value.

  o The labels on the line provide clarity of the values without cluttering the screen with gridlines or prescribing the use of a different style of graph.

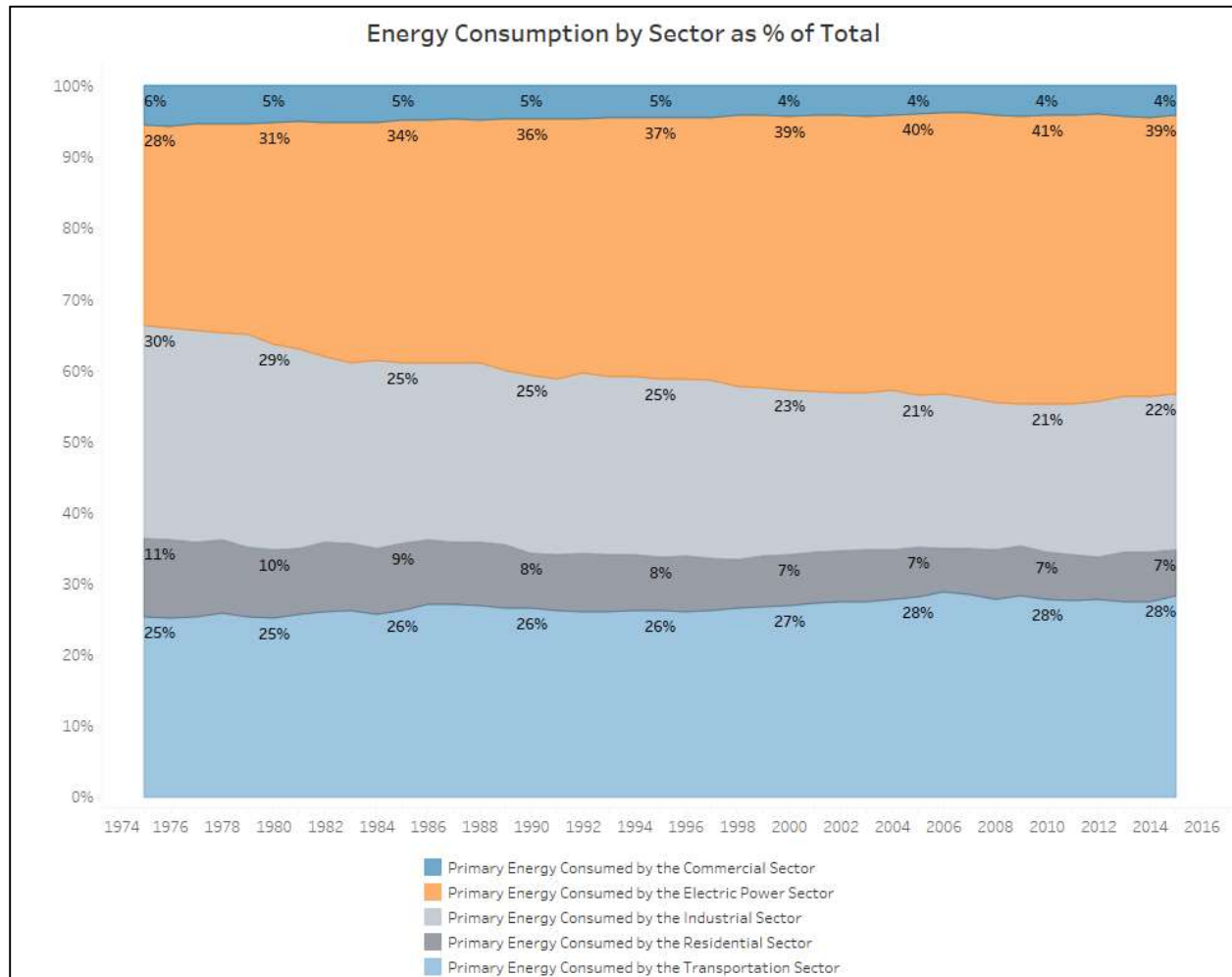*3) Energy Consumption by Sector as a Percentage of Total*

Figure 3 - Energy Consumption by Sector as a Percentage of Total

a) Findings

Fig. 3 provides a striking view of consumption trends in two key consumption sectors: Power Generation and Industrial Consumption. Industrial consumption declined over the time series.

One should not assume that this is simply a result of declining industrial economic activity. It requires further investigation to analyze the impact of technological improvements in the manufacturing process, improved technologies for managing peak demand, and conversion from traditional high-power industries (i.e., steel refineries) to lower power consumers (such as IC chip manufacturing).

b) Intent of Graph

This graph now shows a relationship in declining power consumption and declining demand in the Industrial Sector. It graphically indicates that other sectors remain largely consistent over time, regardless of the degree of reduction in consumption.

c) Design Selections –

- An area chart showing the distribution of total consumed energy per year provides a dramatic view of the decline in Industrial Sector consumption. The use of colors with less saturation reduces competition between the values.

- At this point, colors are assigned to, and remain consistent across consumption sectors (Residential, Industrial, Commercial, Transportation, Electrical Power).

- The Tableau colorblind palette reduces likelihood of color confusion by the reader.

- The selection of the filled area chart provides a cleaner alternative to stacked bars.

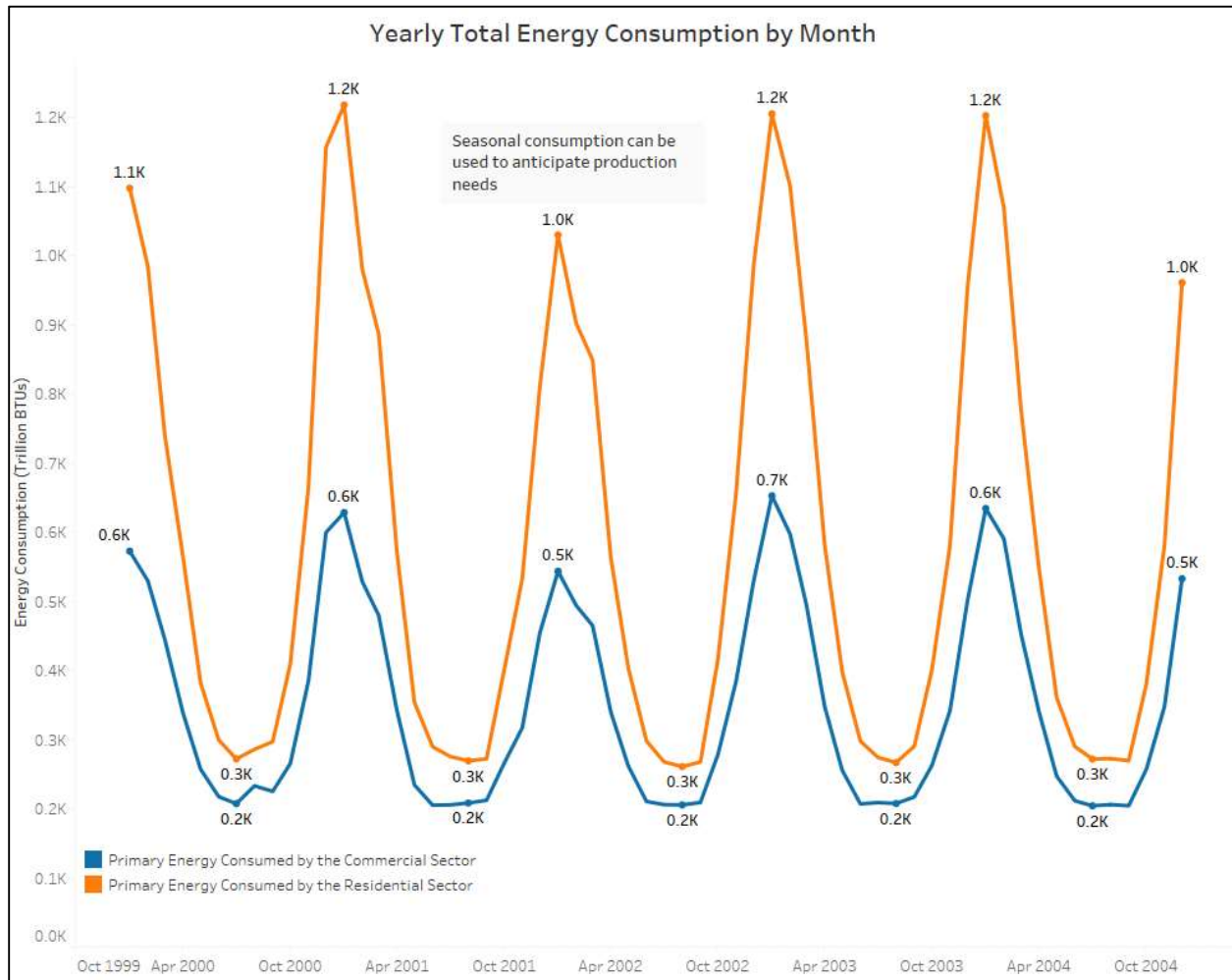*4) Yearly Total Energy Consumption by Month*

Figure 4 0 - Yearly Total Energy Consumption by Month

a) Findings

Fig. 4 pivots to the issue of consumption patterns inherent to the consumption process discussed in previous figures. Commercial and Residential consumption follows the demand for heating and cooling of human-occupied spaces. Cooling costs and extension of lighting hours in the northern hemisphere appear to push a greater demand for power in the winter months compared to lower demand in the summer months.

b) Intent of Graph

This graph moves the reader from persuasion to information. The balance of the report deals with characteristics of consumption and production. This graph eases a change in focus for the causes of changes in consumption to nature of consumption, and finally to the nature of energy generation.

c) Design Selections –

- This document looks at a selection of the graph to demonstrate patterns of consistency. Showing the full time-series would detract from the essential message. However, enough samples were required to establish a pattern that follows from year to year

- The line graph was selected because it provides a mechanism for showing continuous data over a time-series.

- The blue and orange colors are used as the default for the lines and the white background provide contrast and are not likely to confound a color-blind reader.

- Use of the line encoding provides continuity showing the patterns of annual usage.

- Unit information is provided as part of the Y-Axis label (proximity), so it is clear as to its association with the Y-value.

- The labels on the line provide clarity of the values without cluttering the screen with gridlines or prescribing the use of a different style of graph. However, these labels appear on minimum and maximum values for the curve only as displaying all labels would provide too much visual noise to the user and is incidental to display of the pattern.
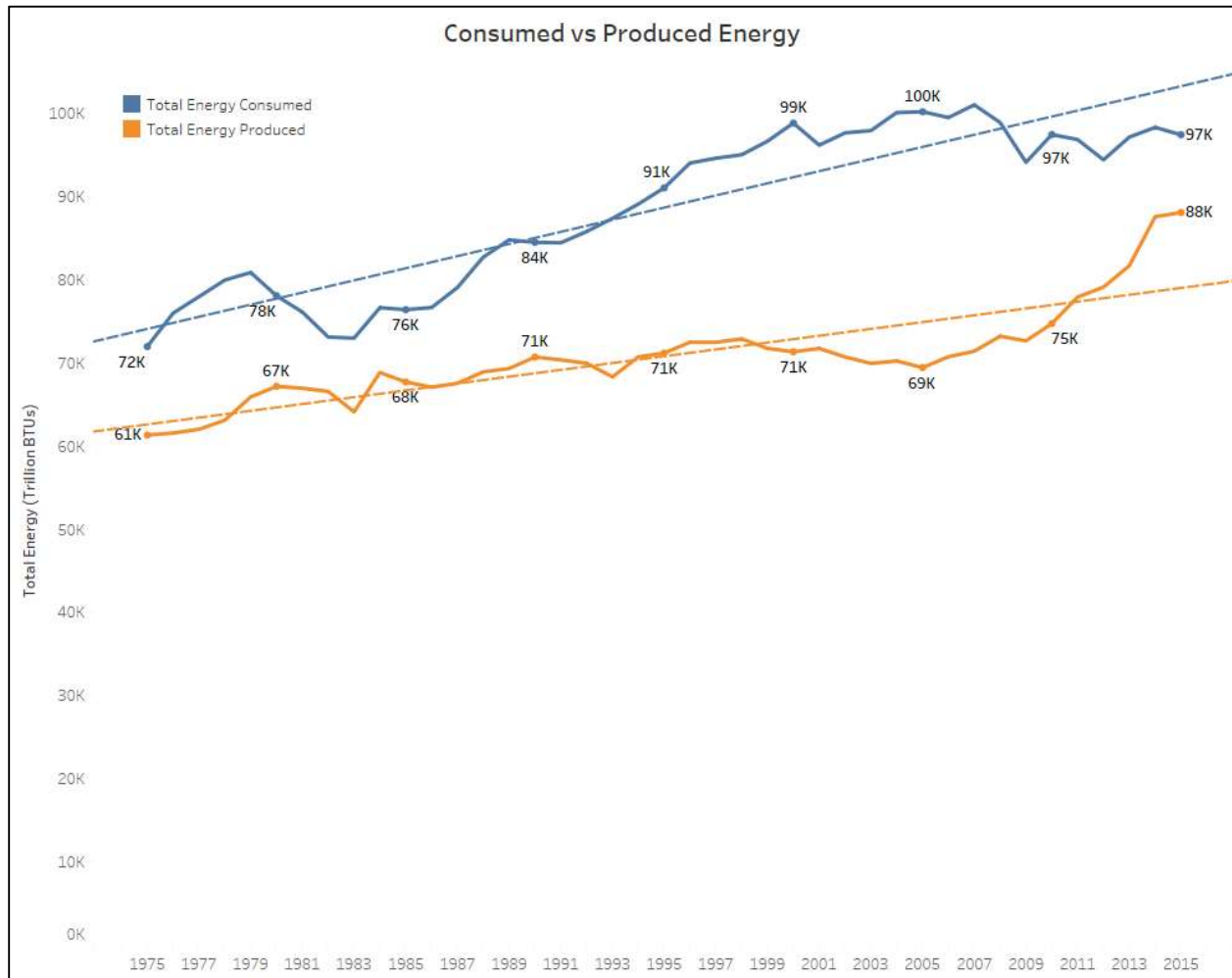
5) Consumed vs. Produced Energy

Figure 5 - Consumed vs. Produced Energy

a) Findings

Fig. 5 shows the relationship between average production and consumption from year to year. Consumption has long outstripped production, although the trend started drifting closer since 2007. The regression lines for each group show the divergence but may require revision based on the constant decline in industrial consumption introduced in Figs. 2 and 3.

b) Intent of Graph

This visualization builds the bridge from production to consumption and introduces the impact of reduced consumption discussed in previous visualizations (Figs. 2 and 3).

c) Design Selections –

- Tableau's dual axis functionality enabled comparison of the different datasets (Production vs. Consumption).

- Trend lines provide a parallel basis of comparison, rather than requiring the reader to try and derive the degree of relationship between the two lines.

- Unit information is provided as part of the Y-Axis label (proximity), so it is clear as to its association with the Y-value.

- The continuation of blue and orange on two category lines continues to support colorblind use.
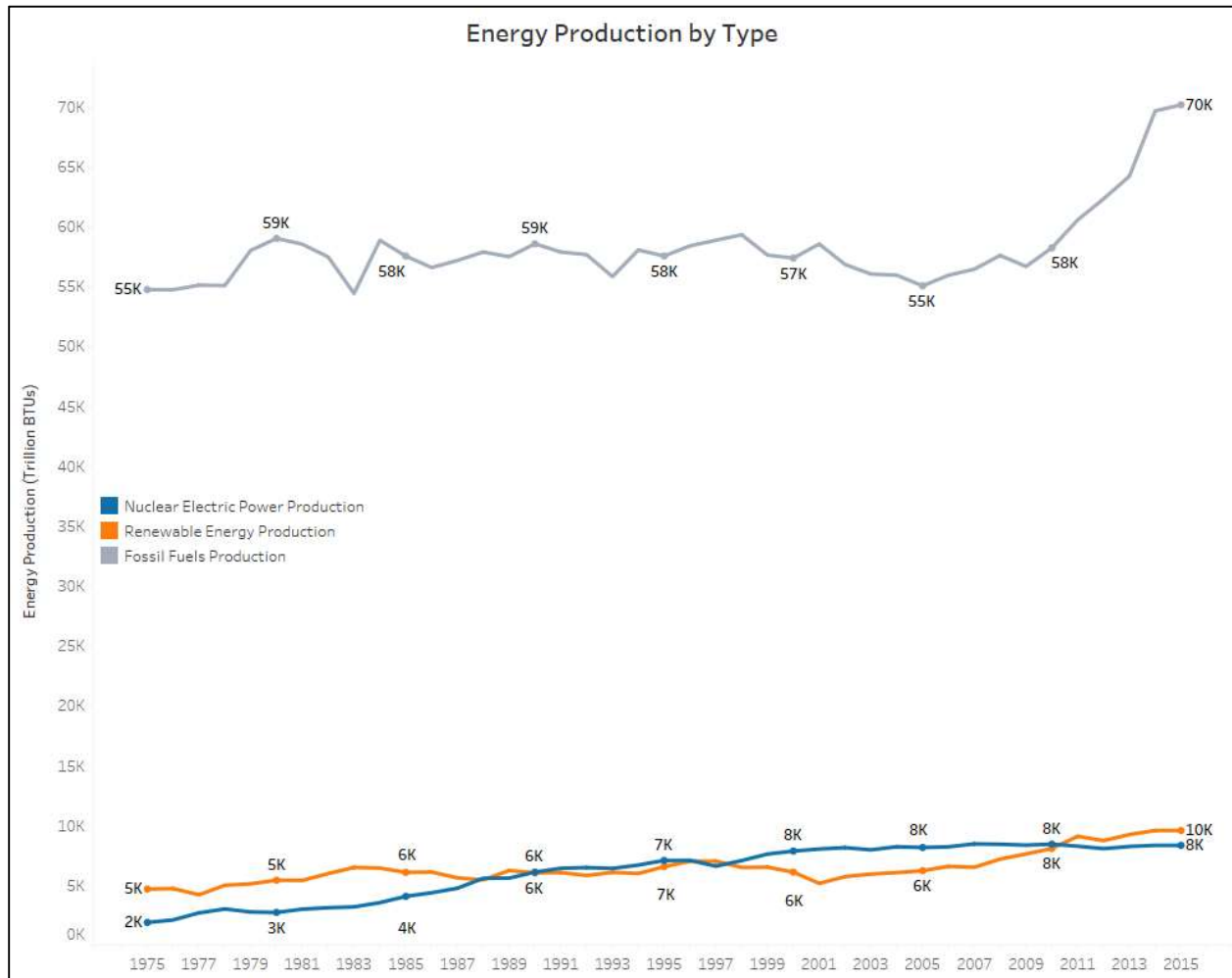
6) Energy Production by Type

Figure 6 - Energy Production by Type

a) Findings

The generation of power in the U.S. consists of three primary categories: Fossil Fuels, Nuclear, and Renewable. Fossil fuel far outstrips the other types. Renewables only recently started to exceed Nuclear Power, excluding a period of higher performance between 1975 and 1989. A rapid uptick in fossil fuel generation far exceeds more moderate increases in renewable and nuclear energy.

b) Intent of Graph

This graph fully completes the pivot to discussing power production with the reader. It also sets the stage for the story to consider dramatic changes in the mix of energies used to generate energy from fossil fuels.

c) Design Selections –

- Lines show time-series information effectively.

- Selective labeling reduces noise and yet provides additional detail allowing for the user to absorb information.

- The graphic series returns to the use of the colorblind palette to avoid confusion for colorblind users.
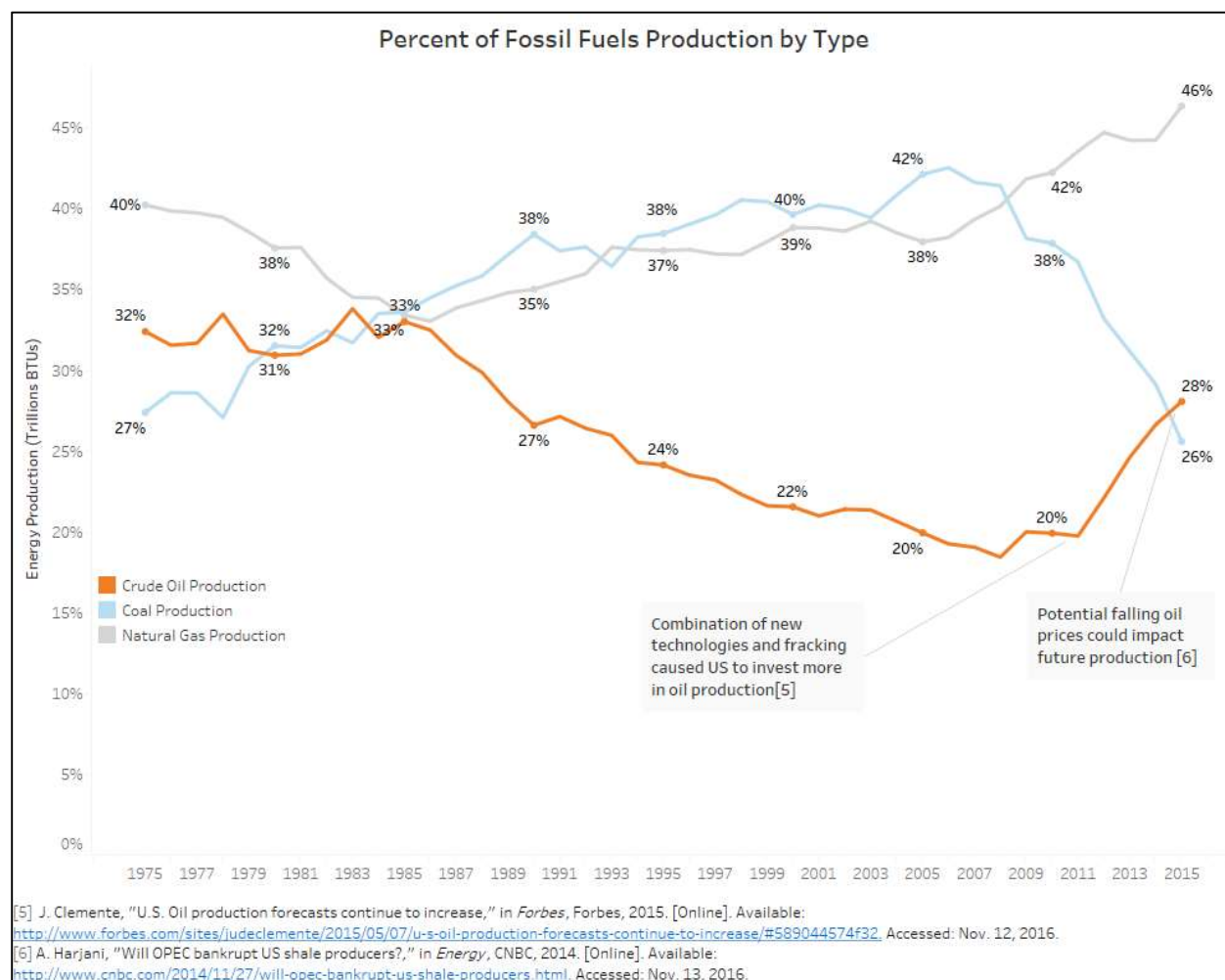
7) Percent of Fossil Fuels Production by Type

Figure 7 - Percent of Fossil Fuels Production by Type

a) Findings

Fossil fuels used for power generation changed substantially over the time series. Coal power generation decreased substantially since 2007 and is at its lowest point since the time series began. Crude oil energy production was at its lowest point in 2008 bust has recovered as crude prices have fallen. Natural gas generation remains on the rise, but historically low oil prices threaten its continued growth.

b) Intent of Graph

This graph shows the reader that fossil fuels now have natural gas centered at the core. Annotations link events that may impact changes in production. This visualization investigates the timing of technological and economic changes on the selection of fossil fuels used in power generation.

c) Design Selections –

- Lines show time-series information.

- High contrast emphasizes Crude Oil power generation as this is critical to the story of technological and economic impact on the oil industry.

- The graphic series continues the use of the colorblind palette to avoid confusion for colorblind users.

- Text boxes and lines provide linkage from the annotations to key events.

- Selective application of labels keeps the lines clear and easy to understand.

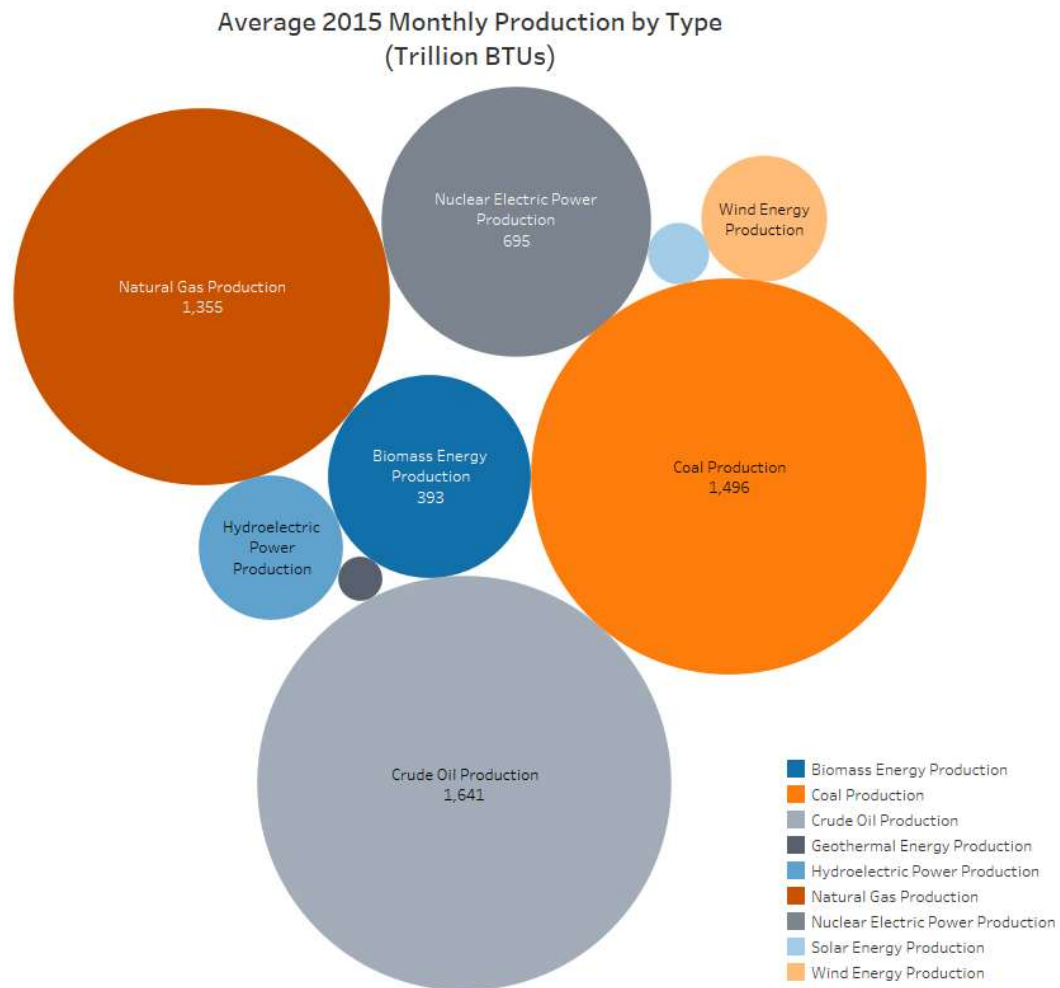*8) Average Monthly Production by Type (Trillion BTUs)*

Average 2015 Monthly Production by Type
(Trillion BTUs)

Natural Gas Production
1,355

Nuclear Electric Power
Production
695

Wind Energy
Production

Biomass Energy
Production
393

Coal Production
1,496

Hydroelectric
Power
Production

Crude Oil Production
1,641

Legend:
- Biomass Energy Production
- Coal Production
- Crude Oil Production
- Geothermal Energy Production
- Hydroelectric Power Production
- Natural Gas Production
- Nuclear Electric Power Production
- Solar Energy Production
- Wind Energy Production

Figure 8 - Average Monthly Production by Type

a) Findings

When considering the generation of energy since the seventies, fossil fuels have been the workhorse. Moreover, while fossil fuel continues to serve as the basis for power generation for the foreseeable future, Natural Gas appears to be growing faster than Crude Oil and Coal.

b) Intent of Graph

This visualization quickly informs the reader of the current state of power generation in the United States. It provides a simple point of closure to encapsulate the information defined over the life of the production.

c) Design Selections –

- Comparative circles show the ranking of energy types.

- Labels on all but the smallest two classes of production reduce unnecessary noise within the categories of power use.

- The graphic series continues the use of the colorblind palette to avoid confusion for colorblind users.

V. REFERENCES

[1] U.S Energy Information Administration, Total Energy, Data, Monthly Energy Review, Data Categories: Energy consumption by sector, Topic: 2.1 Energy consumption by sector CSV file. October 27, 2016. [online] Available: https://www.eia.gov/totalenergy/data/monthly/#appendices Retrieved MER_T02_01.csv file on Friday, November 11, 2016 at 11:25 PM

[2] U.S. Energy Information Administration, Total Energy, Data, Monthly Energy Review, Data Categories: Energy consumption by sector. Topic: C1 Population, U.S. gross domestic product, and U.S. gross output CSV file. October 27, 2016. [Online] Available https://www.eia.gov/totalenergy/data/monthly/#appendices Retrieved MER_TC1.CSV file on Friday, November 11, 2016, at 11:49 PM

[3] U.S. Energy Information Administration, Beta, Total Energy, Data, Change Data Category: 1. Energy Overview, Change Table: Table 1.1 Primary Energy Overview, October 27, 2016. [Online] Available https://www.eia.gov/beta/MER/ Retrieved CSV file on Friday, November 11, 2016, at 11:29 PM

[4] "The Great Recession," in *State of working America*. [Online]. Available: http://stateofworkingamerica.org/great-recession/. Accessed: Nov. 10, 2016.

[5] J. Clemente, "U.S. Oil production forecasts continue to increase," in *Forbes*, Forbes, 2015. [Online]. Available: http://www.forbes.com/sites/judeclemente/2015/05/07/u-s-oil-production-forecasts-continue-to-increase/#589044574f32. Accessed: Nov. 12, 2016.

[6] A. Harjani, "Will OPEC bankrupt US shale producers?," in *Energy*, CNBC, 2014. [Online]. Available: http://www.cnbc.com/2014/11/27/will-opec-bankrupt-us-shale-producers.html. Accessed: Nov. 13, 2016.