Our World in Data Energy

Seyram A. Butame

6/15/23

Our World in Energy

For this week, I wanted to try my hands at creating a visuzal document, and not just one chart. I wanted to put together a single document, that displayed a collection charts, such that the individual document would tell a story. Think in a similar manner to what you see on dashbards, but with static images. Interactivity is sometimes overrated, given that you have to click around all over the place.

Before you do start, be sure to sketch out what you want. Just give you a rough idea of what you want your final out put to look like. I am going for someting that looks like the below.

Figure 1 - Draft plan for the data visualization.

I want to thank Steven Ponce for the code guidance and the inspiration. The visual document he created US energy consumption was similar to what I had in mind (see below). Very nice work.

Figure 2 - Steven Ponce data TidyTuesday visual of US energy consumption.

Libraries

1. `tidytuesdayR` –[1]
2. `tidyverse` – Workhorse suite of packages. Comes with `stringr`, which "provide a powerful and elegant syntax for interpolating strings."[2]
3. `ggtext` – Provides simple Markdown and HTML rendering for ggplot2.[3]
4. `showtext` –[4]
5. `janitor` –[5]
6. `camcorder` –[6]
7. `here` – A Simpler Way to Find Your Files. I have never used this package before but it is great. In particular the `here()` willdinf project files based on the current working directory.[7]

```r
library(pacman)
library(tidyverse)
library(tidytuesdayR)
library(ggtext)
library(showtext)
library(janitor)
library(camcorder)
library(scales)
library(here)
```

Set Up Canvas

I found a nifty little feature in R, where you can set up the canvas that will hold the chart or visual document. Additionally you can record various iterations of your project, that way you look through the progress. This comes via the `camcorder` package and the `gg_record()`. With the gg_record() function, you set up the dimensions of the visual chart or document (breadth, wdith etc.), the format (e.g., png, pdf, jpeg etc.), the directory where it is stored (the `here()` function comes in handy at this point), and set the resolution of the image in dpi.

```r
# FIGURE SIZE/ CANVAS ----
gg_record(
    dir    = here("plots"),
    device = "png",
    width  = 6,
    height = 6,
    units  = "in",
    dpi    = 600)

# TEXT Resolution ----

showtext_opts(dpi = 600) #  Via the showtext package. And sets the resolution options for
```

Load Data

Once we have our canvas set up, we can then read in our data. This week for TidyTuesday we are using the Our World In Data (OWID) Energy data. Lots of different ways to access the tidytuesday data. The `tidytuesdayR` package and its tt_load() function is a good way of accessing the data object via an API. You can then access the specific dataframe from the object via its name. Don't forget to save it. I typically prefer to save them as csv files (i.e.,

comma-separated), as they are light weight and readable in many programs. Take note that I use the `clean_names()` function to read in stored data. This function via the `janitor` package ensures that the resulting names are unique and consist only of the _ character, numbers, and letters. Capitalization preferences can be specified using the case parameter.

The **owid_energy** dataframe contains 129 variables and 21890 observations.

```
## 2. READ IN THE DATA ----

# tuesdata <- tt_load(2023, week = 23)

# owid_energy <- tuesdata$`owid-energy` %>%
#   write_csv(owid_energy, "data/owid_energy.csv")

# write_csv(owid_energy, "data/owid_energy.csv")

owid_energy <- read_csv("data/owid_energy.csv")  %>%
    clean_names()
```

Examine the Data

It is a good to examine a dataframe (i.e., data set) when unfamiliar witht the data. There are a few functions from the dplyr and base R you can use for this process. The `glimpse()` from dplyr is similar to the **describe** command in Stata. However, it provides transposed version of the data with columns running down the page, and data runs across. It's a good way to see all the columns at a go.

The `str()` function or structure function, provides a similar approach to `glimpse()`. But it provides detail in a different way perhaps in a more readable manner. I also use the `view()` function, which will display the data in a spreadsheet format, like you would get with Excel or GoogleSheets. The `head()` from the utils package from base R is also a good option, but it only will show the first five observations and will typically truncate the list of variables. However, the function colnames() will show the list of the columns.

The range() function from base R will show you the min and max values of a particular variable.

```
# ---- EXAMINING THE DATA ----

insp1 <- glimpse(owid_energy)
```

```
Rows: 21,890
```

```
Columns: 129
$ country                      <chr> "Afghanistan", "Afghanist~
$ year                         <dbl> 1900, 1901, 1902, 1903, 1~
$ iso_code                     <chr> "AFG", "AFG", "AFG", "AFG~
$ population                   <dbl> 4832414, 4879685, 4935122~
$ gdp                          <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_cons_change_pct      <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_cons_change_twh      <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_cons_per_capita      <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_consumption          <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_elec_per_capita      <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_electricity          <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_share_elec           <dbl> NA, NA, NA, NA, NA, NA, N~
$ biofuel_share_energy         <dbl> NA, NA, NA, NA, NA, NA, N~
$ carbon_intensity_elec        <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_cons_change_pct         <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_cons_change_twh         <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_cons_per_capita         <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_consumption             <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_elec_per_capita         <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_electricity             <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_prod_change_pct         <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_prod_change_twh         <dbl> NA, 0, 0, 0, 0, 0, 0, 0, ~
$ coal_prod_per_capita         <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0~
$ coal_production              <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0~
$ coal_share_elec              <dbl> NA, NA, NA, NA, NA, NA, N~
$ coal_share_energy            <dbl> NA, NA, NA, NA, NA, NA, N~
$ electricity_demand           <dbl> NA, NA, NA, NA, NA, NA, N~
$ electricity_generation       <dbl> NA, NA, NA, NA, NA, NA, N~
$ electricity_share_energy     <dbl> NA, NA, NA, NA, NA, NA, N~
$ energy_cons_change_pct       <dbl> NA, NA, NA, NA, NA, NA, N~
$ energy_cons_change_twh       <dbl> NA, NA, NA, NA, NA, NA, N~
$ energy_per_capita            <dbl> NA, NA, NA, NA, NA, NA, N~
$ energy_per_gdp               <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_cons_change_pct       <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_cons_change_twh       <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_elec_per_capita       <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_electricity           <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_energy_per_capita     <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_fuel_consumption      <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_share_elec            <dbl> NA, NA, NA, NA, NA, NA, N~
$ fossil_share_energy          <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_cons_change_pct          <dbl> NA, NA, NA, NA, NA, NA, N~
```

```
$ gas_cons_change_twh                <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_consumption                    <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_elec_per_capita                <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_electricity                    <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_energy_per_capita              <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_prod_change_pct                <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_prod_change_twh                <dbl> NA, 0, 0, 0, 0, 0, 0, 0, ~
$ gas_prod_per_capita                <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0~
$ gas_production                     <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0~
$ gas_share_elec                     <dbl> NA, NA, NA, NA, NA, NA, N~
$ gas_share_energy                   <dbl> NA, NA, NA, NA, NA, NA, N~
$ greenhouse_gas_emissions           <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_cons_change_pct              <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_cons_change_twh              <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_consumption                  <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_elec_per_capita              <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_electricity                  <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_energy_per_capita            <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_share_elec                   <dbl> NA, NA, NA, NA, NA, NA, N~
$ hydro_share_energy                 <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_cons_change_pct         <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_cons_change_twh         <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_consumption             <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_elec_per_capita         <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_electricity             <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_energy_per_capita       <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_share_elec              <dbl> NA, NA, NA, NA, NA, NA, N~
$ low_carbon_share_energy            <dbl> NA, NA, NA, NA, NA, NA, N~
$ net_elec_imports                   <dbl> NA, NA, NA, NA, NA, NA, N~
$ net_elec_imports_share_demand      <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_cons_change_pct            <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_cons_change_twh            <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_consumption                <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_elec_per_capita            <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_electricity                <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_energy_per_capita          <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_share_elec                 <dbl> NA, NA, NA, NA, NA, NA, N~
$ nuclear_share_energy               <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_cons_change_pct                <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_cons_change_twh                <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_consumption                    <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_elec_per_capita                <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_electricity                    <dbl> NA, NA, NA, NA, NA, NA, N~
```

```
$ oil_energy_per_capita                          <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_prod_change_pct                            <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_prod_change_twh                            <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_prod_per_capita                            <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_production                                 <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_share_elec                                 <dbl> NA, NA, NA, NA, NA, NA, N~
$ oil_share_energy                               <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewable_consumption                    <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewable_electricity                    <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewable_exc_biofuel_electricity        <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_cons_change_pct               <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_cons_change_twh               <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_elec_per_capita               <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_elec_per_capita_exc_biofuel   <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_energy_per_capita             <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_share_elec                    <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_share_elec_exc_biofuel        <dbl> NA, NA, NA, NA, NA, NA, N~
$ other_renewables_share_energy                  <dbl> NA, NA, NA, NA, NA, NA, N~
$ per_capita_electricity                         <dbl> NA, NA, NA, NA, NA, NA, N~
$ primary_energy_consumption                     <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_cons_change_pct                     <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_cons_change_twh                     <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_consumption                         <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_elec_per_capita                     <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_electricity                         <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_energy_per_capita                   <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_share_elec                          <dbl> NA, NA, NA, NA, NA, NA, N~
$ renewables_share_energy                        <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_cons_change_pct                          <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_cons_change_twh                          <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_consumption                              <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_elec_per_capita                          <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_electricity                              <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_energy_per_capita                        <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_share_elec                               <dbl> NA, NA, NA, NA, NA, NA, N~
$ solar_share_energy                             <dbl> NA, NA, NA, NA, NA, NA, N~
$ wind_cons_change_pct                           <dbl> NA, NA, NA, NA, NA, NA, N~
$ wind_cons_change_twh                           <dbl> NA, NA, NA, NA, NA, NA, N~
$ wind_consumption                               <dbl> NA, NA, NA, NA, NA, NA, N~
$ wind_elec_per_capita                           <dbl> NA, NA, NA, NA, NA, NA, N~
$ wind_electricity                               <dbl> NA, NA, NA, NA, NA, NA, N~
$ wind_energy_per_capita                         <dbl> NA, NA, NA, NA, NA, NA, N~
$ wind_share_elec                                <dbl> NA, NA, NA, NA, NA, NA, N~
```

```
$ wind_share_energy                    <dbl> NA, NA, NA, NA, NA, NA, N~
```

```
insp2 <- owid_energy %>%
  select(iso_code) %>%
  unique() %>%
  arrange(iso_code)

insp3 <- range(owid_energy$year)
```

With the above we first see that our data is in arranged in a long format (i.e.,)

Data Wrangling

1. Hughes, E., Harmon, J., Mock, T. & Community, R. O. L. Access the Weekly TidyTuesday Project Dataset (tidytuesday). *tidytuesdayR* (2019).
2. Wickham, H. *et al.* Welcome to the tidyverse. *Journal of Open Source Software* **4**, 1686 (2019).
3. Wilke, C. O. & Wiernik, B. M. Ggtext: Improved text rendering support for 'ggplot2'. (2022).
4. Qiu, Y. & See file AUTHORS for details., authors/contributors of the included software. Showtext: Using fonts more easily in r graphs. (2022).
5. Firke, S. Janitor: Simple tools for examining and cleaning dirty data. (2023).
6. Hughes, E. Camcorder: Record your plot history. (2022).
7. Müller, K. Here: A simpler way to find your files. (2020).