

A (sun)Shiny Recreation of the New York Times Weather Chart

Amy Crawford, Kiegan Rice, and Katie Will

Abstract

In 1979 The New York Times published what would become a classic visualization of weather data that maps how temperatures compare to historic averages and records over the course of one year. This visualization has changed in several ways over the years, and in 2016 The New York Times published an online, interactive version of this chart, displaying 2015 temperature data for 3,116 world cities. In this project, we scrape historical data from Weather Underground for multiple cities for the years 1945 to 2017. With this data, we recreate the interactive chart published by The New York Times and allow for user interaction via a shiny app. Although we do not include all 3,116 cities, our app includes extra features that allow for users to view the data over the course of many years and choose the features they would like to include on their version of this iconic chart.

1 Introduction

2 Methods

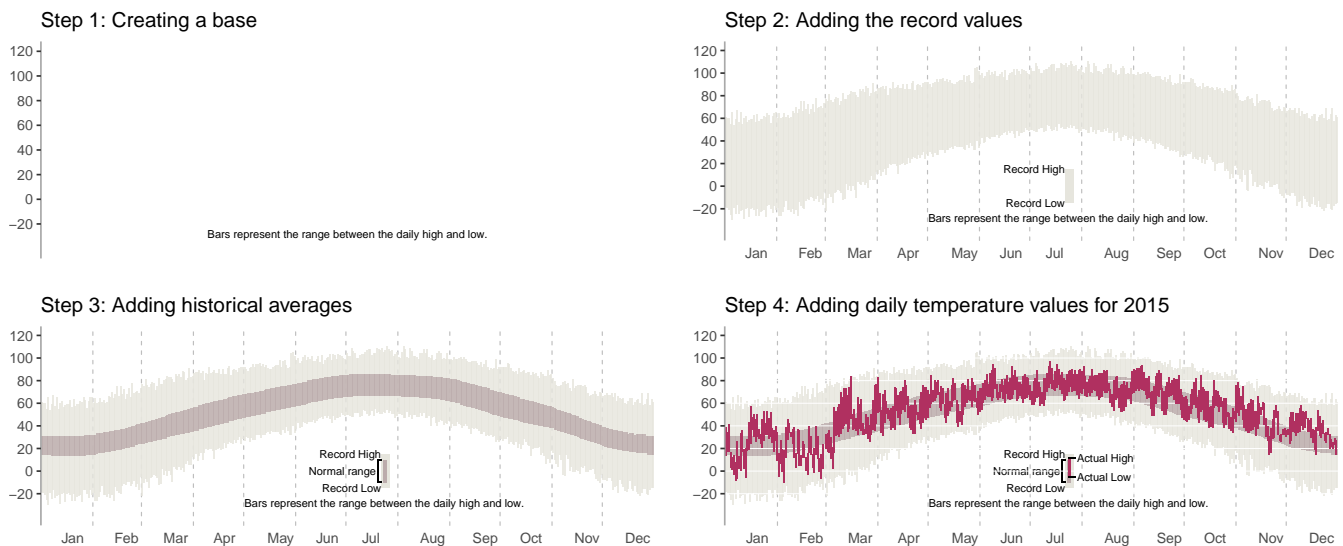


Figure 1: The static visual at each of the individual steps, as layers were created and added to the chart.

2.1 Scraping the Data

2.2 Creating a Static Visual

The next magical step on our rainbow road to finishing 585X was to recreate the static New York Times weather chart. The original published chart

We began by focusing on the city of Des Moines and created our first static visual using the data scraped for this city.

In expectation of needing flexibility of choosing layers in a shiny app, we created the static visual in layers that can be easily peiced together to produce a variety of charts for which the user can choose the detail of the layers to be plotted. The base of the plot included the axes, theme, and legend that would remain consistent regardless of which layers were placed on top. The theme and major gridlines were chosen to match the original chart as closely as possible (with not too much effort though). Each of the dates were plotted for the year 2016 for the base as it includes a leap year and all of our historic averages and records were scraped from the most recent completed year (2016). Thus, regardless of which year the user may choose to plot, leap day averages and records will appear even if a daily temperature is not available. The chart represents three levels of historic weather information and we chose to create each of those layers along with their respective legends in separate stages.

The first stage of the chart was created by graphing bars that spanned from the record low temperature to the record high temperature for each day over the course of one year. We then added a matching legend to the layer that would show up whenever the layer was present on the plot. The next layer was also constructed using bars, these spanned from the average low to the average high temperature for each calendar day, averaged accross all years. This layer also included its own portion of the legend that would plot whenever this layer was present. The final layer that was created included boxes that spanned from the daily high and low temperatures for each calendar day in 2015. This was the top layer and added the final piece of the legend to the plot. Along with this layer we also added faint white y-axis gridlines underneath as a reference point for viewers.

FIGURE 1

2.3 Introducing Shiny Interactivity

2.3.1 Classic Chart Recreation for a Single Year

Although we have a much smaller set of cities in our Shiny app, we wanted to include much of the same interactivity that the New York Times implemented, while also extending to add a few new ways that users could interact with the data. The main update we wanted to include was the ability to change which year the user wanted to have daily temperatures for. Although the original graphic allows for a lot of cool interpretability, the ability to compare across years is a huge advantage when investigating how weather trends have changed. Although we have daily temperature data ranging all the way back to 1945, the set of cities we are working with are inconsistent with their start dates. However, all cities in our data have daily temperatures beyond 1950, so we chose to make this element of the interactivity available for 1950 to 2016 only. Since the chart spans a full year, we do not include 2017 in the ‘Single Year’ tab of our Shiny app. Users can interact with data for the current year in the other tab, explained in section **something**.

As already discussed in the scraping section, we have a limited scope of cities. However, the interactivity with city is essentially the same as the New York Times chart, with the exception of the super cool moving globe thing. Users can choose a midwestern city from a drop-down menu that includes all available cities (that the app creators have scraped up to this point).

Users of the app can select which year they want to look at by moving a slider input across a variety of years. When users select a year, they are selecting which year they would like to see for the top layer. It is important to note that the record and historical averages layers will not change for any given city, no matter which year is selected. These are record values that are current through March 2017, either averaged with data from the current year (for average data), or updated when record data happens.

The most innovative update to the interactive version of this chart (or at least we think so) is that the user may choose which layers of the plot they would like to view together. A user can look at any of the three layers individually, or any combination of the three layers. This allows users to see the chart in all its glory, or just examine a single aspect of the data for that given city and year. The base layer and gridlines remain constant, and as users check boxes to select which layers they want to examine, the layers are added or removed from this base plot accordingly. No matter which order the boxes are checked in, layers will always be added in the order they are meant to be in the original chart (i.e., records on the bottom, then averages, and finally the daily temperatures on top). This interactivity was made possible by a reactive function that creates 7 plots of the possible layer combinations for the city-year pair, and returns the one that matches the boxes that the user has checked.

2.3.2 Historic Spaghetti Monster Plot for a Single City

3 Results (maybe call this Outcomes?)

4 Discussion

5 Future Work

- Fix data inconsistencies
- Add in magic turning world
- Optimizing code
- Adding in ability to switch to celsius (?)
- Adding more cities
- Add some sort of plotly interactivity/denote record years