portfolio 4

Jianzhuo Liu

2022-05-08

```r
library("dplyr")
```

```
## Warning: package 'dplyr' was built under R version 4.0.5
```

```r
library("tidyverse")
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
## Warning: package 'tibble' was built under R version 4.0.5
```

```
## Warning: package 'tidyr' was built under R version 4.0.5
```

```r
library("ggplot2")
library("ggridges")
library("readr")
library("patchwork")
```

```
## Warning: package 'patchwork' was built under R version 4.0.5
```

```r
weather <- read_csv("D:/GREmfzl/STAT 610 HW/weatherAUS.csv")
```

```
## Warning: 153782 parsing failures.
## row         col              expected actual                                file
## 6050 Evaporation 1/0/T/F/TRUE/FALSE   12   'D:/GREmfzl/STAT 610 HW/weatherAUS.csv'
## 6050 Sunshine    1/0/T/F/TRUE/FALSE   12.3 'D:/GREmfzl/STAT 610 HW/weatherAUS.csv'
## 6051 Evaporation 1/0/T/F/TRUE/FALSE   14.8 'D:/GREmfzl/STAT 610 HW/weatherAUS.csv'
## 6051 Sunshine    1/0/T/F/TRUE/FALSE   13   'D:/GREmfzl/STAT 610 HW/weatherAUS.csv'
## 6052 Evaporation 1/0/T/F/TRUE/FALSE   12.6 'D:/GREmfzl/STAT 610 HW/weatherAUS.csv'
## .... ........... .................. ...... ......................................
## See problems(...) for more details.
```

```r
weather <- weather[!is.na(weather$RainToday),]
weather <- weather[!is.na(weather$RainTomorrow),]
weather_analyze <- weather%>%
mutate(datemodify = lubridate::yday(Date))%>%
group_by(Location,datemodify)%>%
filter(Location == "Albury" || Location == "Uluru")
p<-list()
```

```
p[["ribbon"]]<-ggplot(weather_analyze)+geom_ribbon(aes(Date,ymin=MinTemp, ymax=MaxTemp,fill=Location),al
scale_x_continuous(expand = c(0, 0)) +
scale_y_continuous(expand = c(0, 0, 0.1, 0)) +
scale_size_continuous(range = c(0.1, 4))
p[["rain"]] <- ggplot(weather_analyze) +
geom_point(aes(Date, WindGustSpeed, col = Location, size = Rainfall),alpha=0.8) +
scale_x_continuous(expand = c(0, 0)) +
scale_y_continuous(expand = c(0, 0, 0.1, 0)) +
scale_size_continuous(range = c(0.1, 4))
p[["range"]] <- ggplot(weather_analyze) +
geom_histogram(
aes(MaxTemp-MinTemp, fill = Location),
alpha = 0.8, position = "identity"
) +
scale_y_continuous(expand = c(0, 0, 0.1, 0))
p[["range"]] <- p[["range"]] +
labs(x = "Temp Range", title = "Temp Diff")
p[["ribbon"]] <- p[["ribbon"]] +
labs(x = "Date", y = "Temperature", title = "Temp Trend")
p[["rain"]] <- p[["rain"]] +
labs(x = NULL, title = "WindSpeed & RainFall") +
scale_color_discrete(guide = "none")
p[["range"]] + (p[["ribbon"]] / p[["rain"]]) +
plot_layout(guides = "collect", widths = c(1, 3))
```

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 31 rows containing missing values (geom_point).
```
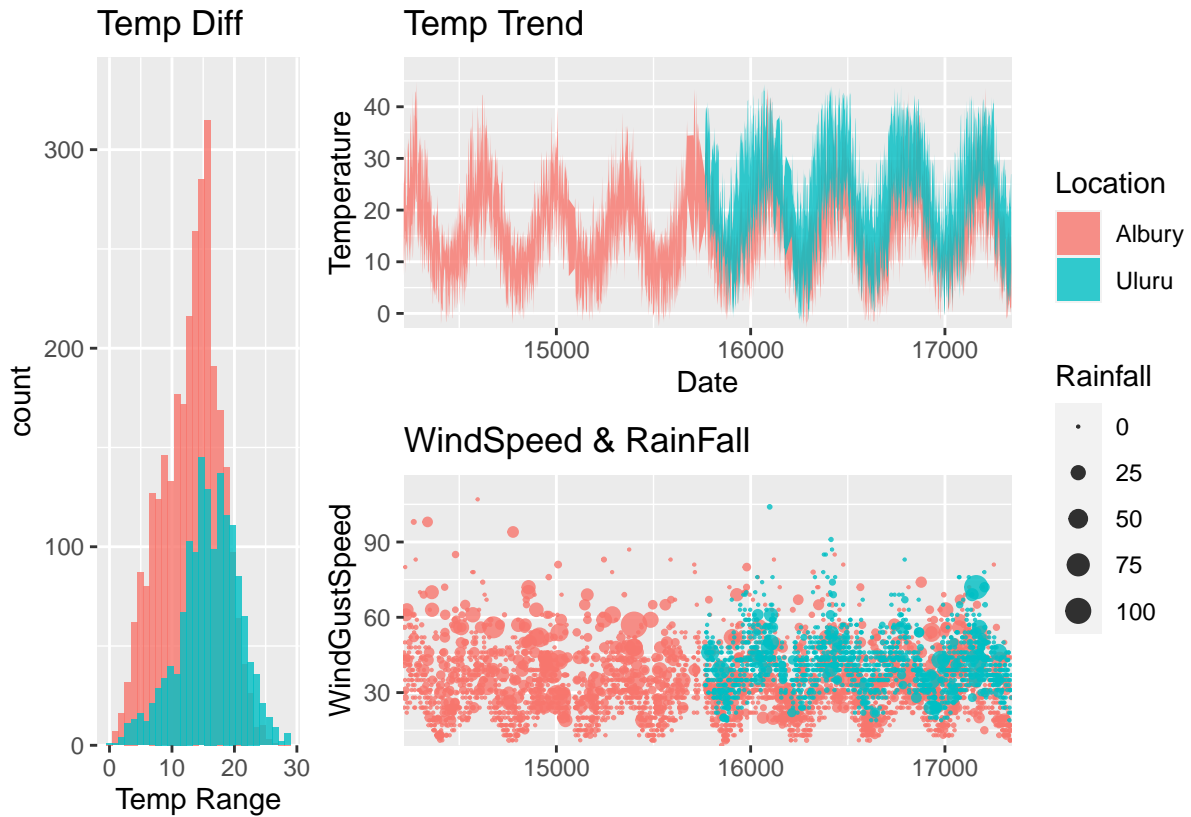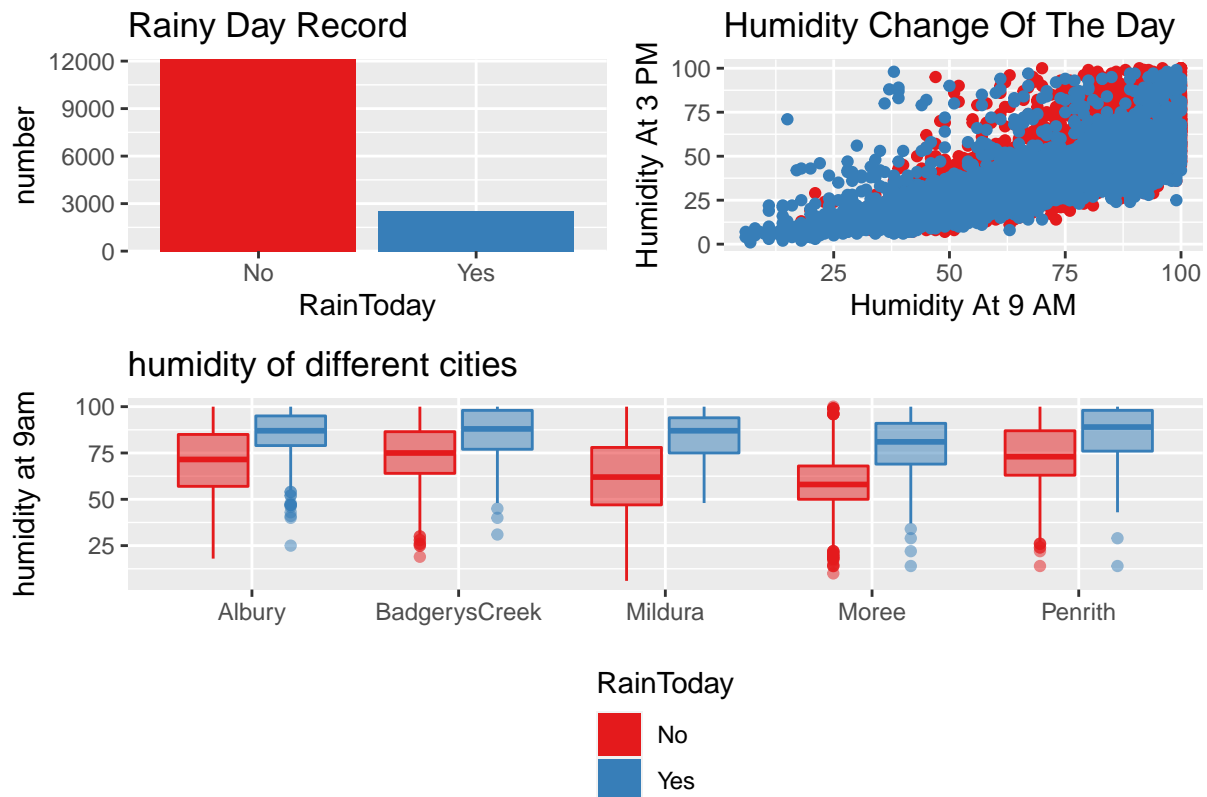
```
weather_analyze <- weather%>%
mutate(datemodify = lubridate::yday(Date))%>%
group_by(Location,datemodify)%>%
  filter(Location == "Albury" || Location == "Mildura")
weather_analyze2 <- weather%>%
mutate(datemodify = lubridate::yday(Date))%>%
group_by(Location,datemodify)%>%
filter(Location == "Albury" || Location == "BadgerysCreek" || Location == "Mildura" || Location == "More
p<-list()
p[["bar"]] <- ggplot(count(weather_analyze2, RainToday)) +
geom_bar(aes(RainToday, n, fill = RainToday), stat = "identity") +
scale_y_continuous(expand = c(0, 0)) +
scale_fill_brewer(palette = "Set1") +
labs(y = "number")+ggtitle("Rainy Day Record")
p[["scatter"]] <- ggplot(weather_analyze) +
geom_point(aes(Humidity9am, Humidity3pm, col = Location)) +
scale_color_brewer(palette = "Set1") +
theme(legend.position = "none") +
labs(x = "Humidity At 9 AM", y = "Humidity At 3 PM")+
ggtitle("Humidity Change Of The Day")
p[["box"]] <- ggplot(weather_analyze2) +
geom_boxplot(aes(Location, Humidity9am, col = RainToday, fill = RainToday), alpha = 0.5) +
scale_color_brewer(palette = "Set1") +
scale_fill_brewer(palette = "Set1") +
theme(legend.position = "none") +
labs(y = "humidity at 9am", x = NULL)+ggtitle("humidity of different cities")
```

```
(p[["bar"]] + p[["scatter"]])/ p[["box"]] + plot_layout(guides = "collect") &
plot_annotation(theme = theme(legend.position = "bottom"))
```

```
## Warning: Removed 13 rows containing missing values (geom_point).
```

```
## Warning: Removed 80 rows containing non-finite values (stat_boxplot).
```



```
library("tidyverse")
library("tsibble")
```

```
## Warning: package 'tsibble' was built under R version 4.0.5
```

```
library("lubridate")
library("feasts")
```

```
## Warning: package 'feasts' was built under R version 4.0.5
```

```
## Warning: package 'fabletools' was built under R version 4.0.5
```

```
library("scales")
weather_analyze3 <- weather%>%
mutate(datemodify = lubridate::yday(Date))%>%
```

```
group_by(Location,datemodify)%>%
mutate(year=year(Date))%>%
filter((Location == "Uluru") & (year=="2016" ||year == "2015"||year=="2013"||year=="2017"||year=="2014")
weather_analyze3  <- weather_analyze3 [!duplicated(weather_analyze3 ),]
weather_modi_tsibble <- weather_analyze3 %>%
  as_tsibble(index = Date,key = Location)
weather_modi_tsibble <- fill_gaps(weather_modi_tsibble, .full = TRUE)
gg_season(weather_modi_tsibble, y = MaxTemp, period = "year") +
  facet_grid(Location~.) +
  ggtitle("Max Temperature in Uluru from 2013 to 2017") +
  labs(colour = "Year") +
  xlab("Month in the Year") +
  ylab("Max Temp") +
  theme(axis.text=element_text(size=10), plot.title = element_text(size = 15, face = "bold", hjust = 0.5
```


Max Temperature in Uluru from 2013 to 2017