A. R. Onojeghuo

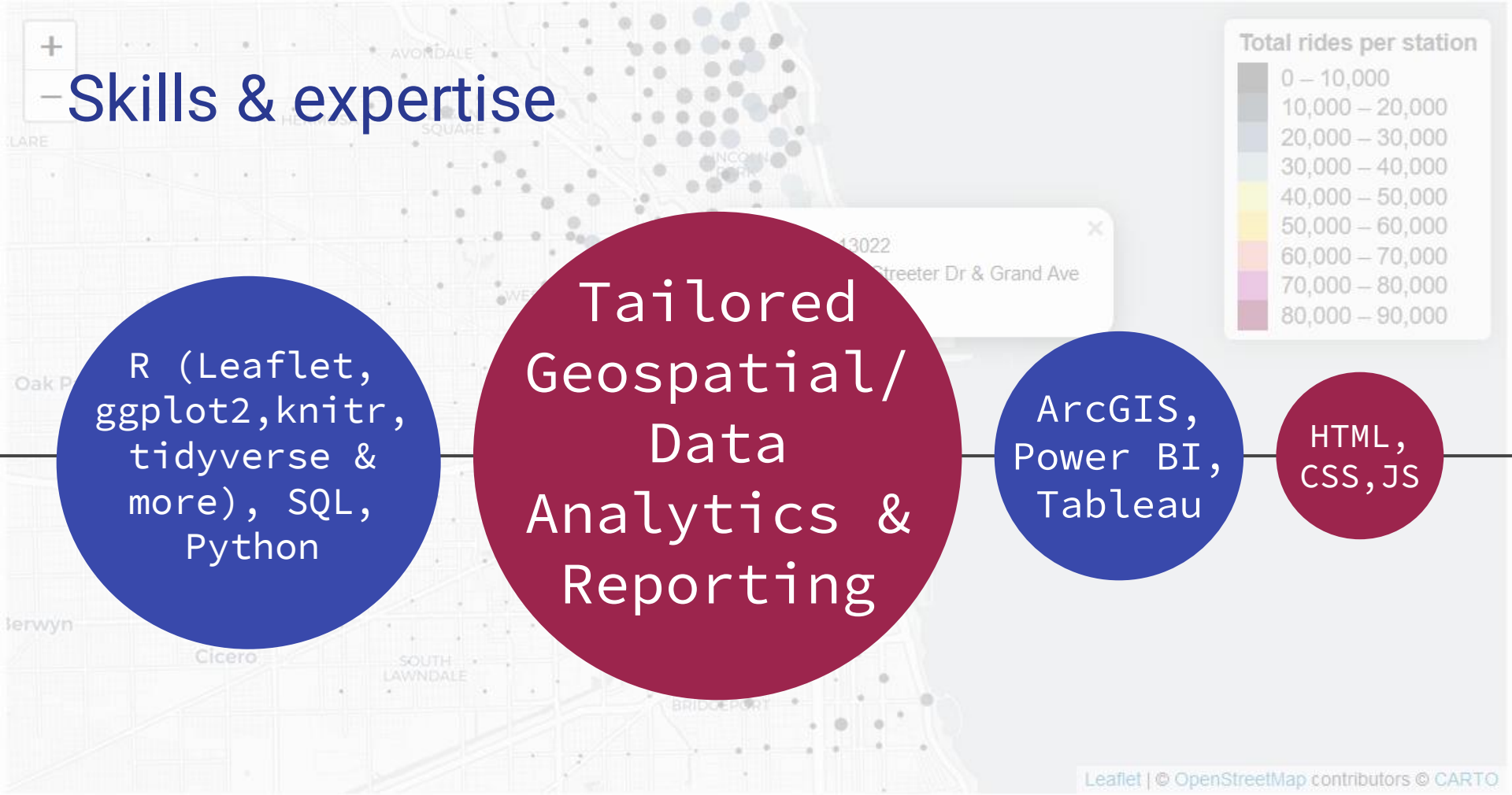
PhD, Geog (Sci)

Data Scientist/GIS & Remote Sensing Specialist

About me

I’m a result-oriented and experienced Data Science, GIS & Remote Sensing Specialist with over 10 years of experience. I have hands-on experience using Excel, Tableau, SQL, R (ggplot2, tidyverse, Rmarkdown, leaflet, knitr), Power BI, ArcGIS, and more for data analytics and visualization. I have a good command of English and relay technical results in simple terms

Skills & expertise



R (Leaflet, ggplot2,knitr, tidyverse & more), SQL, Python

Tailored Geospatial/ Data Analytics & Reporting

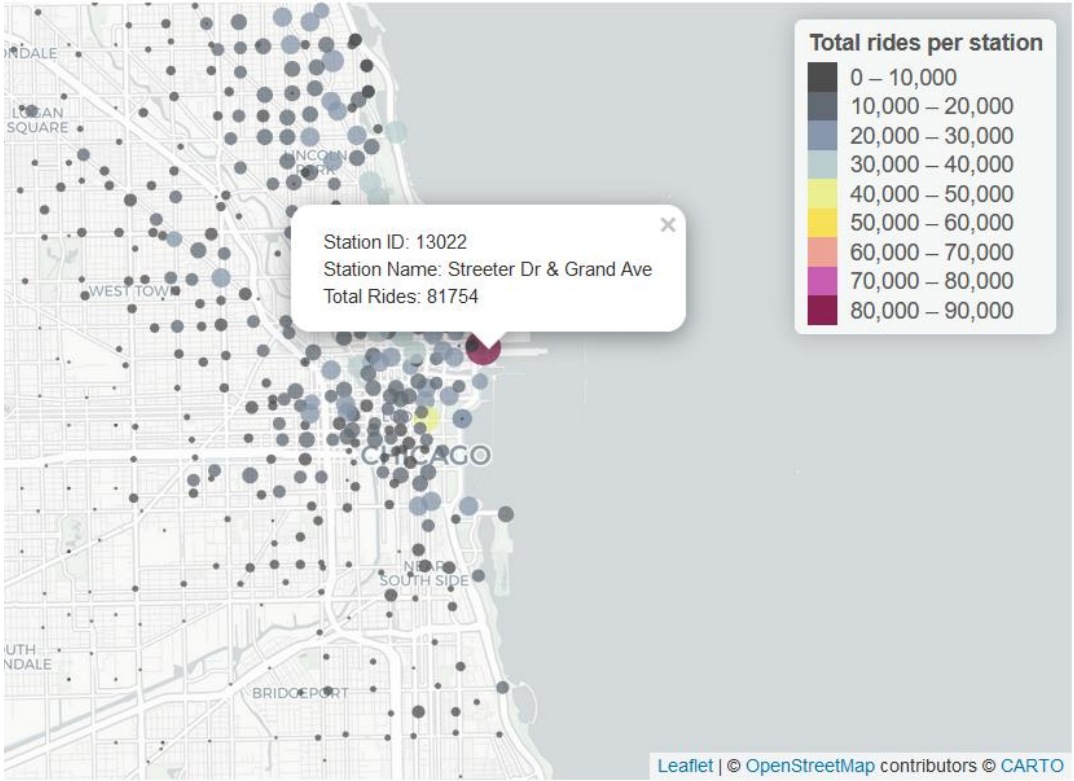
ArcGIS, Power BI, Tableau

HTML, CSS,JS

Portfolio samples

Bike-share trend analysis using R

- Analysis to determine bike-share trends in the Chicago area for Cyclistic (Apr 2021 - Mar 2022).
- Data was downloaded from https://divvy-tripdata.s3.amazonaws.com under a license agreement for educational purposes.
- Data download, unzipping, merging and cleaning was done in R.



Bike-share trend analysis (contd.): Sample R code for descriptive analysis

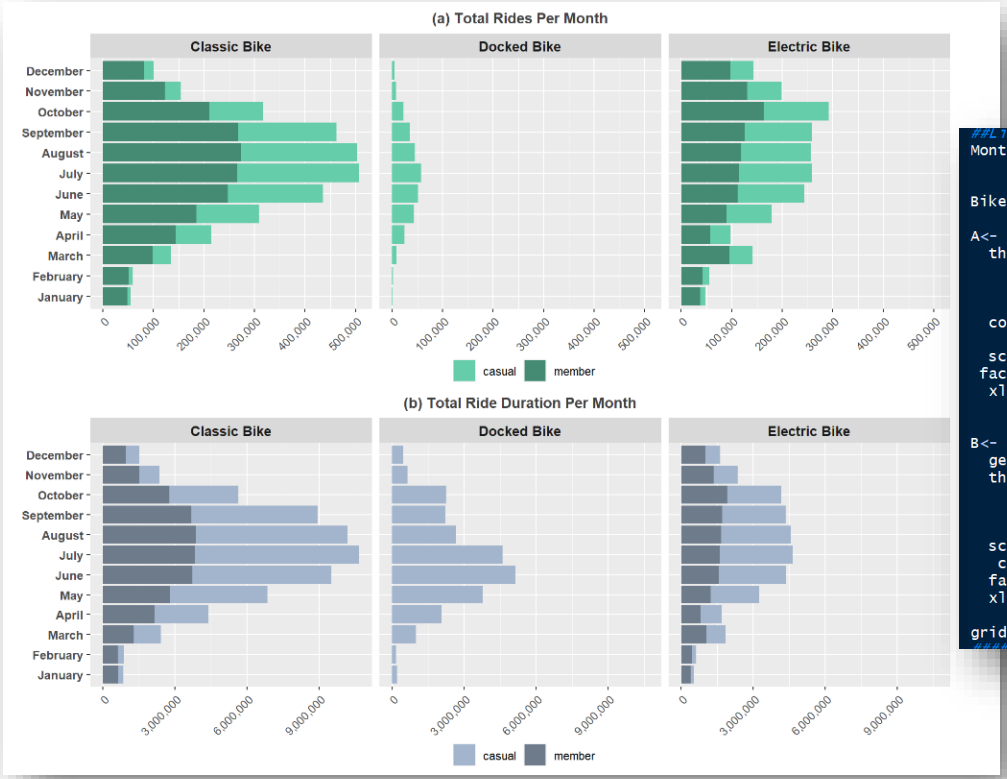
```r
SA <-read_csv("Cyclistic202104_202203.csv" )
##Delete duplicate ride ids and fill in missing values using other similar rows
S1 <- SA %>%
  group_by(start_station_id,start_station_name) %>% fill(start_lat, start_lng) %>% ungroup %>%
  group_by(end_station_id,end_station_name) %>% fill(end_lat, end_lng) %>% ungroup %>%
  group_by(start_lat, start_lng)%>% fill(start_station_id,start_station_name) %>% ungroup %>%
  group_by(end_lat, end_lng) %>% fill(end_station_id,end_station_name) %>% ungroup

##distinct(ride_id, .keep_all= TRUE) %>%

#%>% group_by()
S1$RideLength <- difftime(S1$ended_at, S1$started_at,units="mins")
##Extract day of the week. 1=Sunday and 7 is Saturday
S1$day_of_the_week<-wday(S1$started_at)
S1$Hour <-hour((S1$started_at))
S1$Month<-month((S1$started_at))
S1$Wd2 <-wday(S1$started_at,label=T)
```

```r
##List all files
L1 <- list.files(pattern ="20" )
##Create empty list to store column names
XLlist <-list()
##read files and extract column names
for (i in L1) {
 A1 <- read.csv(file = i, header = T)
 A2 <-  as.data.frame(colnames(A1))
XLlist[[i]] <- A2
}

XList2 <- list.cbind(XLlist)
L2 <- L1 %>%
  lapply(read_csv) %>%    # Store all files in list
  bind_rows # Combine data sets into one data set
write.csv(L2,"Cyclistic202104_202203.csv")
##Install Census API key if needed
```

Bike-share trend analysis (contd.): Most popular bike-share station

```r
##To find the mode of the months of incidents
Mode = function(x){
  ta = table(x)
  tam = max(ta)
  if (all(ta == tam))
    mod = NA
  else
    if(is.numeric(x))
      mod = as.numeric(names(ta)[ta == tam])
  else
    mod = names(ta)[ta == tam]
  return(mod)
}

MinMode <-  function(x){
  ta = table(x)
  tam = min(ta)
  if (all(ta == tam))
    mod = NA
  else
    if(is.numeric(x))
      mod = as.numeric(names(ta)[ta == tam])
  else
    mod = names(ta)[ta == tam]
  return(mod)
}

S2 <- S1 %>%  filter(!start_lat >=45) %>% group_by(member_casual) %>%
  summarize(Most_popular_takeoff_Station_id=Mode(start_station_id),
            Most_popular_takeoff_Station=Mode(start_station_name),
           Most_popular_destination= Mode(end_station_name),Mosttravelled_Hour=Mode(Hour),
            Weekday_most_travel= Mode(day_of_the_week),Mean_rideLength=mean(RideLength),
            Max_ridelength= as_hms(max(RideLength)),Month_Most_Travel=Mode(Month))
write.csv(S2,"Cyclistic_summary2.csv")
```

Bike-share trend analysis (contd.): Monthly bike rides by members and casual riders



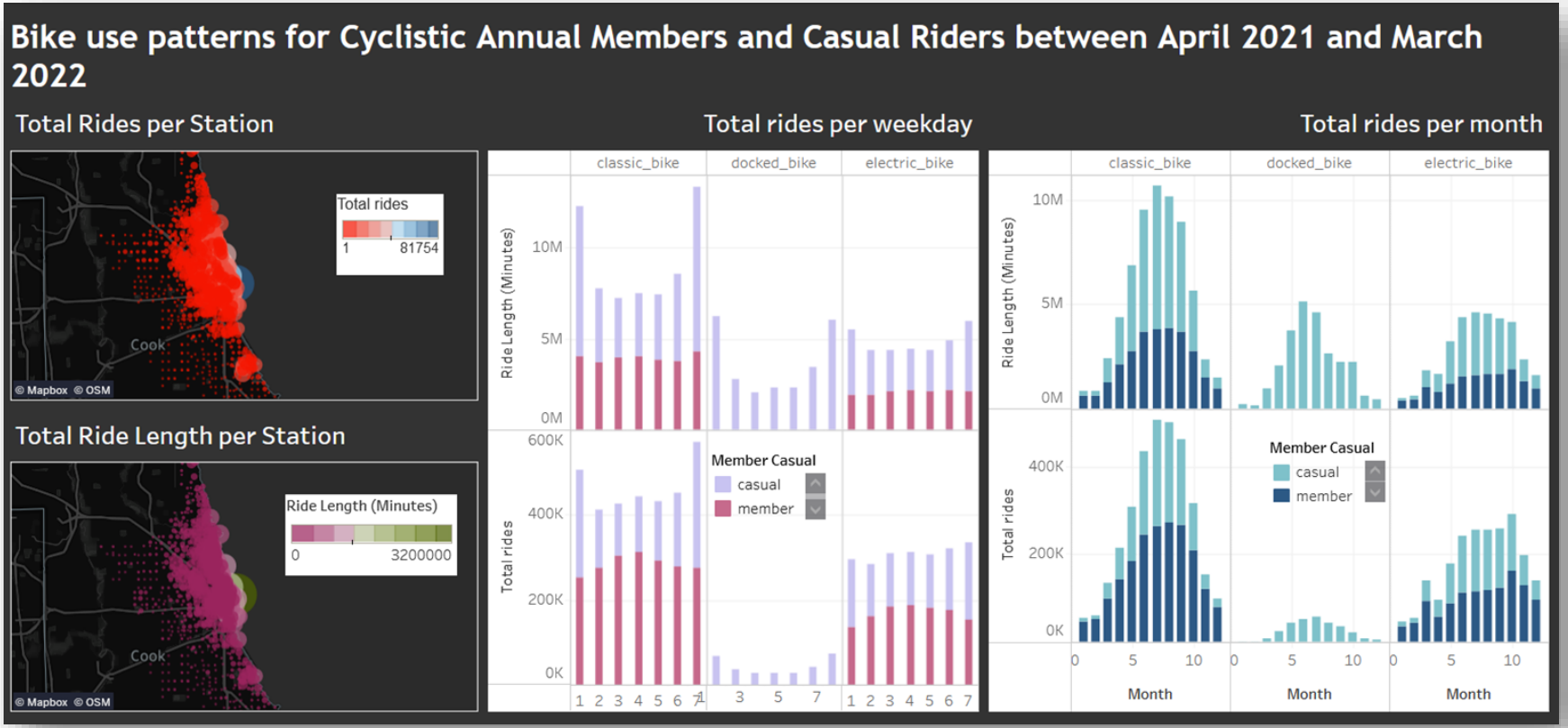

```
##List Months
Months<- c("January","February","March","April","May","June",
           "July","August","September","October","November","December")

Bikes<-c("classic_bike"="Classic Bike","docked_bike"="Docked Bike",
         "electric_bike"="Electric Bike")
A<- ggplot(CYM, aes(fill=member_casual, y=Total_rides, x=as.factor(Month))) + theme_gray() +
  theme(legend.position = "bottom",axis.text.x = element_text(angle = 45,hjust=1,vjust=1),
        axis.text.y = element_text(face="bold", size=9),
        strip.text = element_text(face="bold", size=10),legend.margin=margin(-20, 0, 0, 0),
        plot.title = element_text(color="gray25",size=12,hjust = 0.5,face="bold"))+
  coord_flip()+ggtitle("(a) Total Rides Per Month")+
    geom_bar(position="stack", stat="identity")+
  scale_fill_manual("",values=c("aquamarine3","aquamarine4")) +
 facet_wrap(~ rideable_type,labeller = as_labeller(Bikes))+ ylab("")+
  xlab("")+scale_x_discrete(labels= Months)+ scale_y_continuous(labels = comma)


B<- ggplot(CYM, aes(fill=member_casual, y=Ride_Length, x=as.factor(Month))) +
  geom_bar(position="stack", stat="identity") +theme_gray () +
  theme(legend.position = "bottom",axis.text.x = element_text(angle = 45,hjust=1,vjust=1),
        axis.text.y = element_text(face="bold", size=9),
        strip.text = element_text(face="bold", size=10),legend.margin=margin(-20, 0, 0, 0),
        plot.title = element_text(color="gray25",size=12,hjust = 0.5,face="bold"))+
  scale_fill_manual("",values=c("lightsteelblue3","lightsteelblue4")) +
   coord_flip() +ggtitle("(b) Total Ride Duration Per Month")+
  facet_wrap(~ rideable_type,labeller = as_labeller(Bikes)) + ylab("")+
  xlab("")+scale_x_discrete(labels= Months)+ scale_y_continuous(labels = comma)

grid.arrange(A,B,ncol=1)
#####################################################################
```

Bike-share trend analysis (contd.): Weekly bike rides by members and casual riders

- More casual riders used the bike-share service in summer months and on weekends (suggesting leisure & tourism related travel)
- Bike-share was mostly from Station 13022 (Streeter Dr & Grand Ave) for casual riders and from Kingsbury St & Kinzie St for annual members

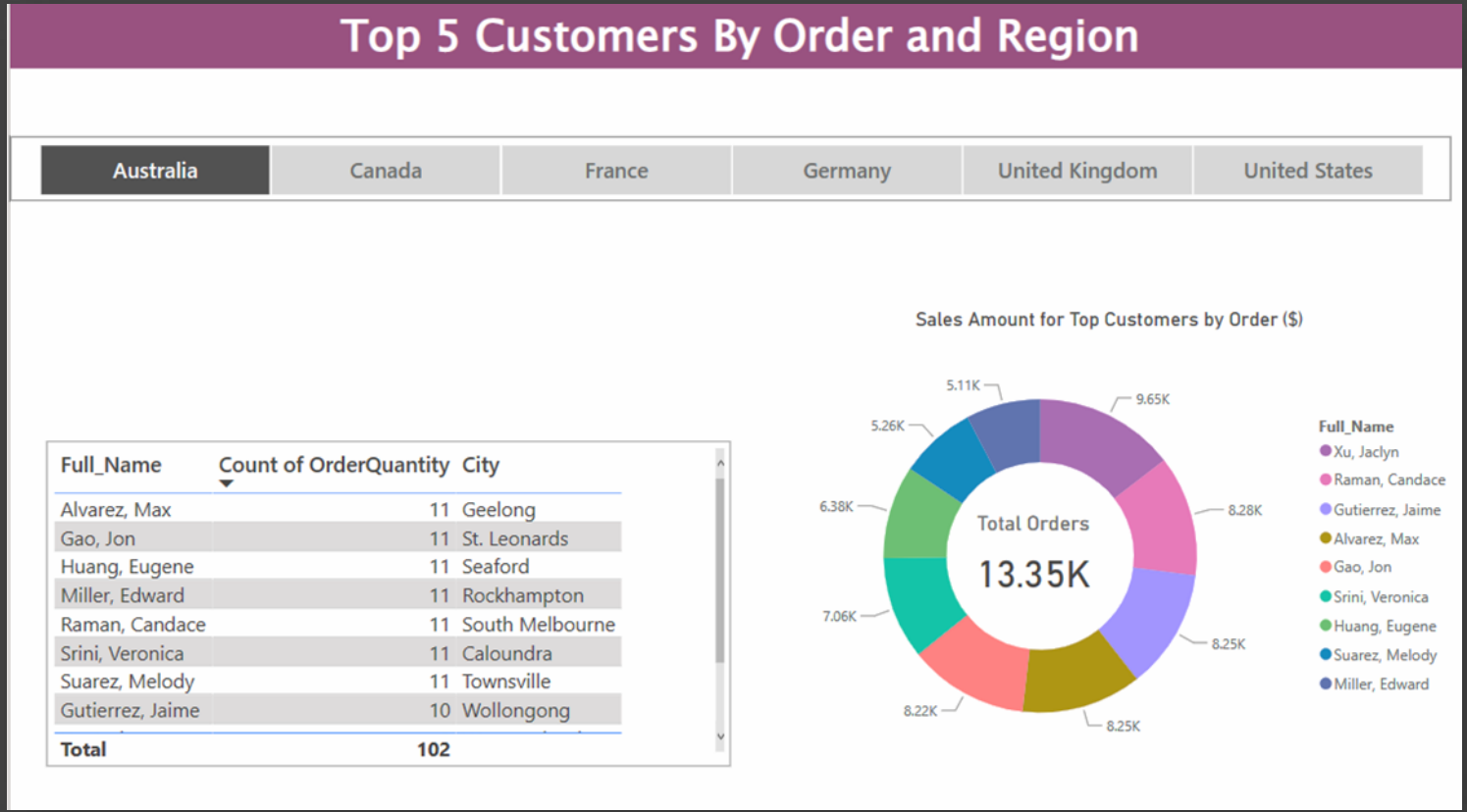
Full report available at: https://rpubs.com/aronojeghuo/Google-Cyclistic
Code repository: https://github.com/aonojeghuo/Google-Data-Analytics-Capstone-Project/blob/main/Capstone_RKnit%20report%20from%20Markdown4.Rmd

Bike-share trend analysis (contd.): Visualization using Tableau Public



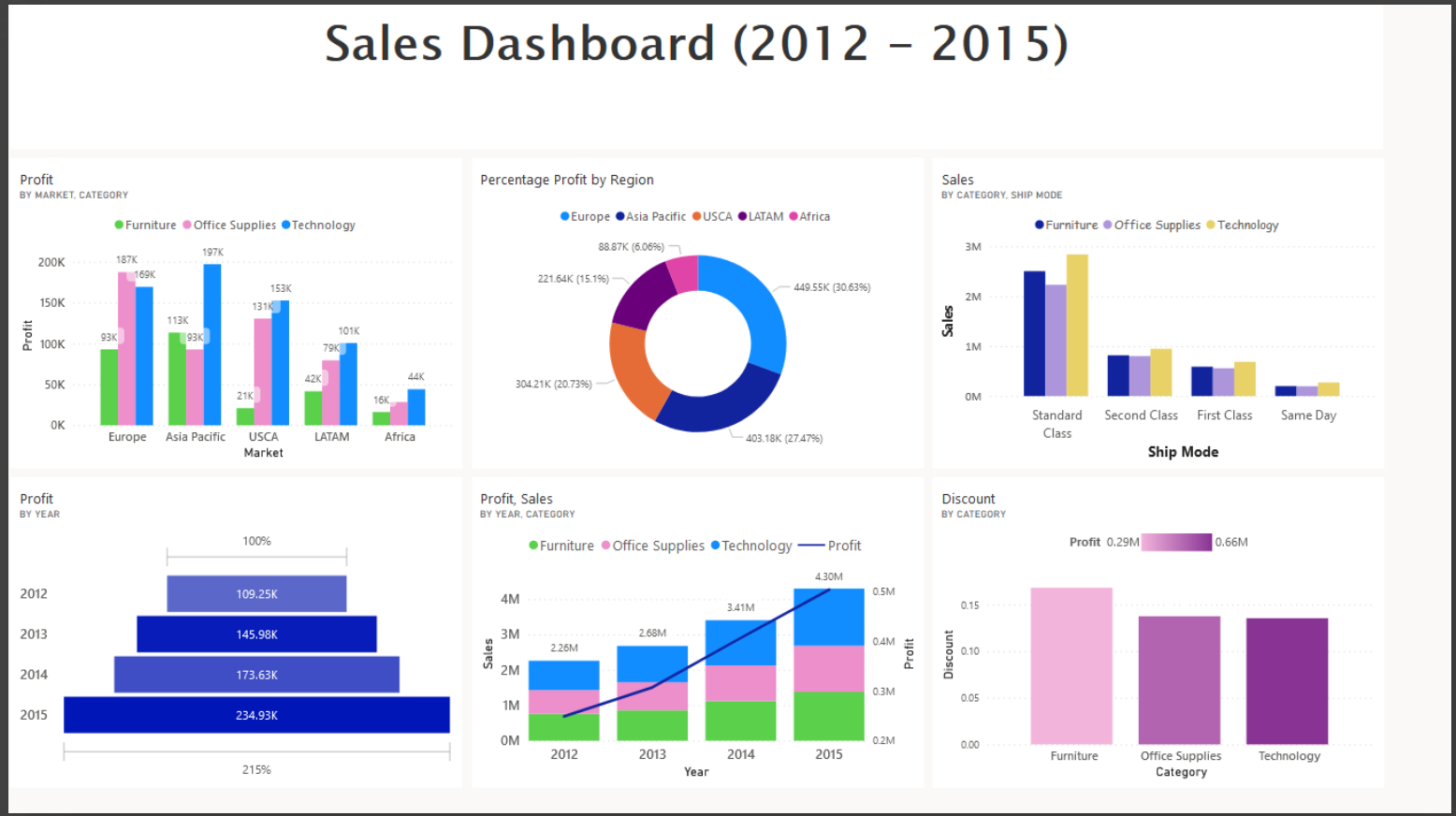
https://public.tableau.com/app/profile/ajoke.onojeghuo/viz/CyclisticDataAnalysis_16527444370630/Dashboard1

Sales analysis using Power BI Desktop



Power BI sheet showing top customers by value of total purchases per country. The slider allows the client shift between countries.

Sales analysis using Power BI (contd.): Power BI Service App



Power BI App showing a dashboard of Sales. It features a 4-year profit trend and percentage profit by region.

Career highlights

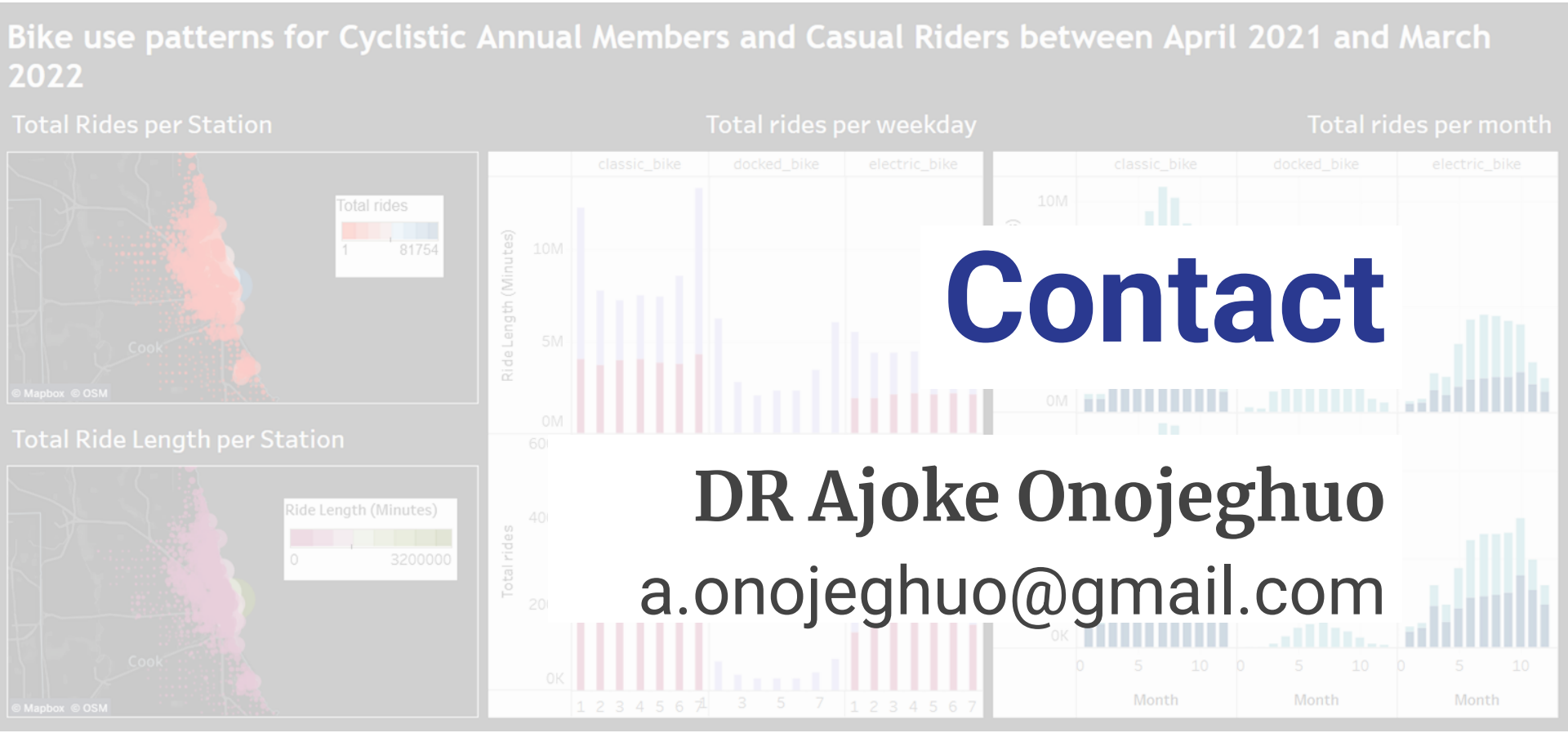
Data Scientist/GIS & Remote Sensing Specialist
Solstice Environmental Management, Edmonton
November 2019 - February 2022

Data Scientist/GIS & Remote Sensing Specialist
Jolexy Environmental Services, Edmonton
September 2012 - Till Date

GIS Analyst/ Data Scientist
PLACE Research Lab, School of Public Health, University of Alberta, Edmonton
March 2018 - October 2019

Sample Publication

- Onojeghuo, A. O., **Onojeghuo, A. R.**, Cotton, M., Potter, J., & Jones, B. (2021). Wetland mapping with multi-temporal sentinel-1 &-2 imagery (2017–2020) and LiDAR data in the grassland natural region of Alberta. GIScience & Remote Sensing, 1-23.
- Rasul, A., Ibrahim, S. A., **Onojeghuo, A. R**., & Balzter, H. (2020). A trend analysis of leaf area index and land surface temperature and their relationship from global to local scale. *Land*, *9*(10), 388.
- **Onojeghuo, A. R.**, Nykiforuk, C. I., Belon, A. P., & Hewes, J. (2019). Behavioral mapping of children's physical activities and social behaviors in an indoor preschool facility: methodological challenges in revealing the influence of space in play. International journal of health geographics, 18(1), 1-16.
- Onojeghuo, A.O and **Onojeghuo, A.R.** (2017). Object-based habitat mapping using very high spatial resolution multispectral and hyperspectral imagery with LiDAR data. International Journal of Applied Earth Observation and Geoinformation, Vol. 59, 79–91 (Publisher: Elsevier).


Bike use patterns for Cyclistic Annual Members and Casual Riders between April 2021 and March 2022
Total Rides per Station
Total rides
1
81754
Cook
© Mapbox © OSM
Total Ride Length per Station
Ride Length (Minutes)
0
3200000
Cook
© Mapbox © OSM
Total rides per weekday
classic_bike
docked_bike
electric_bike
Ride Length (Minutes)
10M
5M
0M
Total rides
60
40
20
0K
1 2 3 4 5 6 7
Total rides per month
classic_bike
docked_bike
electric_bike
10M
0M
0K
0 5 10
Month
Contact
DR Ajoke Onojeghuo
a.onojeghuo@gmail.com