Sprint-3

Date	17 Nov2022
TeamID	PNT2022TMID19258
Project Name	Project –SignswithSmartConnectivityforBetterRoadSafety

- ApplicationPackages
- Main
- Slides
- Textcenter_analysis
- Task_Rmd

ApplicationPackages:

```
package

com.example.myhp.accidentprevention;importandroid.a

pp.Application;
importandroid.test.ApplicationTestCase;
/***
<ahref="http://d.android.com/tools/testing/testing_android.html">TestingFund>
*/
publicclassApplicationTestextendsApplicationTestCase<Application>
  {
      publicApplicationTest()
{
      super(Application.class);
   }
}
```

MAIN:

\#Manipulatingdata{#data}

Thissectionisanintroductiontomanipulatingdatasetsusingthe`dplyr`package. Asoutlinedintheprevioussection,`dplyr`and `ggplot2`arepartofthe`tidyverse`,whichaimstoprovideauser-friendlyframeworkfordatascience[@grolemund_data_2016].

ExperienceofteachingRoverthepast fewyearssuggeststhat manypeoplefinditeasiertogetgoingwithdatadrivenresearchiftheylearnthe'tidy'workflowpresented inthissection.
However,ifyoudonotlikethisstyleofRcodeoryouaresimplycurious,weencourageyoutotryalternativeapproachesforachievingthesimilarresultsusingbaseR[@rcoreteam_language_2020]^[
Runthecommand`help.start()`toseearesources introducingbaseR,and [Chapter6onlistsanddataframes](https://cran.r-project.org/doc/manuals/r-release/R-intro.html#Lists-and-data-frames)in[AnIntroductiontoR](https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf)inparticularforanintroductiontodatamanipulationwithbaseR.
]
,the`data.table`Rpackage[@R-data.table]orotherlanguagessuchas[Python](https://www.python.org/)or[Julia](https://julialang.org/).
Ifyoujustwanttoget goingwithprocessingdata,the`tidyverse`isasolidandpopularstartingpoint.

<!--Todo: addnewparthere?-->

Beforedivingintothe`tidyverse`, itisworthre-cappingwherewehavegottoso faraswehavecoveredalotofground.
Section\@ref(basics)introducedR'sbasicsyntax;Section\@ref(rstudio)showedhowtousetheSource EditorandotherfeaturesofRStudiotosupportdatascience;andSection \@ref(pkgs)introducedtheconceptandpracticalitiesofRpackages,withreferenceto`stats19`,`ggplot2`and
`dplyr`.

Inthissection,wewillstartwithablankslate.
InSection\@ref(basics)welearnedthatinRhavinga'cleardesk'meansan*emptyglobalenvironment*.
Thiscanbeachievedbyrunningthefollowingcommand,whichremovesthe`list()`ofallobjectsreturnedbythefunction `ls()`:

```{r}
rm(list=ls())
```

##tibbles

AlthoughthedataprocessingtechniquesinRarecapableofhandlinglargedatasets,suchasthe `crashes_2019`objectthatwecreatedintheprevioussection,representing100k+casualties,itmakessensetostartsmall.
Let'sstart byre-creatingthe`crashes`datasetfromSection\@ref(basics),butthistimeusingthe `tidyverse``tibble()`function.Thisisthe`tidyverse`equivalentofbaseR's`data.frame`.
`tibble`objectscanbecreated,afterloadingthe`tidyverse`,asfollows:

```{r,message=FALSE}l
ibrary(tidyverse)crashes
=tibble(
  casualty_type=c("pedestrian","cyclist","cat"),cas
  ualty_age                                    =seq(from
  =20,to=60,by=20),vehicle_type=c("car","bus",
  "tank"),
  dark=c(TRUE,FALSE,TRUE)
)
```

Inthepreviouscodechunk, wepassedfourvectorobjectsas*namedarguments*tothe`tibble`function,resultingincolumnssuchas `casualty_type`.
A`tibble`isjust afancywayofrepresenting`data.frame`objects,preferredby`tidyverse`usersandoptimisedfordatascience.
Ithasafewsensibledefaultsandadvantagescomparedwiththe`data.frame`,oneofwhichcanbeseenby printinga`tibble`:

```{r}class(crashes
)crashes
```

Notethe`<chr>`,`<dbl>`or`<lgl>`textbeloweachcolumn,providingaquickindicationoftheclassofeachvariable- this isnotprovidedwhenusing`data.frame`.

## filter()andselect()rowsandcolumns

Intheprevioussection,webrieflyintroducedthepackage`dplyr`, whichprovidesanalternativetobaseRformanipulatingobjects.`dplyr`providesdifferent,andsomewouldarguesimpler,approaches for subsetting rowsandcolumnsthan baseR.
`dplyr`operationsforsubsettingrows(withthefunction`filter()`)and columns(withthefunction`select()`)aredemonstrated below.Herewecanalsoseetheuseofthepipeoperator `%>%`totakethedatasetandapplythefunctiontothat dataset.

```{r}
crashes %>%filter(casualty_age >50)#
filtersrowscrashes%>%select(casualty_type)#selectjustone
column
```

Itshouldbeclearwhathappened:`filter()`returnsonlyrowsthatmatchthecriteriainthefunctioncall,onlyobservationswitha`casualty_age`greaterthan50 inthiscase.
Likewise,`select()`returnsdataobjectsthatincludeonlycolumnsnamedinsidethefunctioncall, `casualty_type`inthiscase.

Togainagreaterunderstandingofthefunctions,typeandrunthefollowingcommands,whichalsoillustratehowthe`%>%`canbeusedmorethanoncetomanipulatedata(moreonthissoon):

```{r}
```

crashes_darkness=crashes%>%filter(dark)crashes_a=crashes%>%select(contains("a"))crashes_darkness_a = crashes%>%filter(dark)%>%
  select(contains("a"))

Canyouguesswhatthedimensionsoftheresultingobjectswillbe? Writedownyourguessesforthenumberofrowsandnumberofcolumnsthatthenewobjects, `crashes_darkness`to`crashes_darkness_a`,havebeforerunningthefollowingcommandstofindout. Thisalsodemonstratesthehandyfunction`dim()`,shortfordimension(resultsnotshown):^[
Notethatthenumberofrowsisreportedbeforethenumberofcolumns. ThisisafeatureofR:rowsarealsospecificiefiedfirstwhensubsettingusingthesquarebracketsincommandssuchas`crashes[1,2:3]`.
]

```{r,eval=FALSE}dim(crashes)
dim(crashes_darkness)
?contains#gethelponcontains()tohelpguesstheoutputofthenextlinedim(crashes_a)
dim(crashes_darkness_a)
```

Lookatthehelppagesassociatedwith`filter()`,`select()`andtherelatedfunction`slice()`asfollowsand tryrunningtheexamplesthatyouwillfindatthebottomofthehelppagesforeachtogainagreaterunderstanding(noteyoucanusethe`package::function`notationtogethelponfunctionsalso):

```{r,eval=FALSE}
?dplyr::filter
?dplyr::select
?dplyr::slice
```

##Orderingandselectingthe'topn'

Other usefulpipe-friendlyfunctionsare `arrange()`and`top_n()`. `arrange()` canbe usedtosortdata.Within.the`arrage()`function,optionalargumentscanbeusedtodefinetheorderin whichitissorted.`top_n()`simplyselectsthetop'n'numberofrowsinyourdataframe. Wecanusethesefunctionstoarrangedatasetsandtakethetopmost‘n’values,asfollows:

```{r}crashes
%>%
```

  arrange(vehicle_type)crashe
s%>%
  top_n(n=1,wt=casualty_age)

<!--##Longandwidedata-->

##Summarise

Apowerfultwo-
functioncombinationis`group_by()`and`summarise()`.Usedtogether,theycanpr
ovide*groupedsummaries*ofdatasets.
Intheexamplebelow,wefindthemeanageofcasualtiesindarkandlight conditions.

```{r}crashes
%>%
  group_by(dark)%>%
  summarise(mean_age=mean(casualty_age))
```

Theexampleaboveshowsapowerfulfeatureofthesepipelines.Manyoperationscanbe'chained'togeth
er,whilstkeepingreadabilitywithsubsequentcommandsstackedbelowearlieroperations.Thecombin
ationof`group_by()`and`summarise()`canbeveryusefulinpreparingdataforvisualisationwitha`ggpl
ot2`function.
Anotherusefulfeatureofthe`tidyverse`fromauserperspectiveistheautocompletionofcolumnnames
midpipe.
Ifyouhavenotnoticedthisalready,
youcantestitbytypingthefollowing,puttingyourcursorjustbeforethe`)`and pressing `Tab`:

```{r,eval=FALSE}
crashes%>%select(ca)#pressTabwhenyourcursorisjustafterthea
```

Youshouldsee`casualty_age`and`casualty_type`popupasoptionsthatcanbeselectedbypressing`Up
`and `Down`.
Thismaynotseemlikemuch,but whenanalysinglargedatasetswithdozensofvariables,it
canbeagodsend.

Ratherthanprovidingacomprehensiveintroductiontothe`tidyverse`suiteofpackages,thissections
houldhaveofferedenoughtogetstarted
withusingitforroadsafetydataanalysis.Forfurtherinformation,checkoutup-to-
dateonlinecoursesfromrespectedorganisationslike[DataCarpentry](https://datacarpentry.org/R-
ecology-
lesson/index.html)andthefreeonline[books](https://bookdown.org/)suchas[RforDataScience]((
https://r4ds.had.co.nz/))[@grolemund_data_2016].

##Tidyverseexercises

1. Use`dplyr`tofilterrowsinwhich`casualty_age`islessthan18, andthen28.
2. Usethe`arrange`functiontosortthe`crashes`objectindescendingorderofage(**Hint:**seethe`?arrange` helppage).
3. Readthehelppageof`dplyr::mutate()`.What doesthefunctiondo?
4. Usethemutatefunctiontocreateanewvariable,`birth_year`, inthe`crashes`data.framewhichisdefinedasthecurrentyearminustheirage.
5. **Bonus:**Usethe`%>%`operatorto filtertheoutputfromtheprevious exercisesothatonlyobservationswith`birth_year`after1969arereturned.

~~~~
```{rdplyr,eval=FALSE,echo=FALSE}
#answerscrashes
%>%
  arrange(desc(casualty_age))
crashes%>%filter(casualty_age>21)cras
hes%>%
  mutate(birth_year=2019-
  casualty_age)%>%filter(birth_year>1969)
```
~~~~

Slides:

~~~~
---
title:"RoadSafety(andtransport)ResearchwithR"
#subtitle:'`remojifont::emoji("bike")`<br/>ForEnglandandWales'subtitle:'`r
emojifont::emoji("rocket")`<br/>RACFoundation,DataDriven'author:"Robi
nLovelace"
date:'2020'outpu
t:
  xaringan::moon_reader:
    #css:["default","its.css"]
    #chakra:libs/remark-
    latest.min.jslib_dir:libs
    nature:
      highlightStyle:githubhighlightLines:
      true

#bibliography:
#-../vignettes/ref.bib
#-../vignettes/ref_training.bib
---

```{rsetup,include=FALSE,eval=FALSE}
#getcitations
refs=RefManageR::ReadZotero(group="418217",.params=list(collection="JFR868KJ",limit=
100))
refs_df=as.data.frame(refs)
#View(refs_df)
~~~~

~~~
#citr::insert_citation(bib_file="vignettes/refs_training.bib")RefManageR::WriteBib(refs,
"refs.bib")
#citr::tidy_bib_file(rmd_file="vignettes/pct_training.Rmd",messy_bibliography="vignett
es/refs_training.bib")
options(htmltools.dir.version=FALSE)knitr::o
pts_chunk$set(message=FALSE)library(RefM
anageR)BibOptions(check.entries=FALSE,
        bib.style="authoryear",c
        ite.style='alphabetic',sty
        le="markdown",first.ini
        ts=FALSE,hyperlink=F
        ALSE,dashed=FALSE)
my_bib=refs
```
~~~

~~~
```{r,eval=FALSE,echo=FALSE,engine='bash'}
#publishresultsonline
cp-Rvcode/rrsrr-slides*~/saferactive/site/static/slides/cp-
Rvcode/libs~/saferactive/site/static/slides/
cd~/saferactive/sitegit
add-A
gitstatus
gitcommit-
am'Updateslides'gitpush
cd-
```
~~~

#Slide/linkshttps://itsleeds.githu

b.io/rrsrr/https://bookdown.org/

https://www.pct.bike/

---

background-image:url(https://media.giphy.com/media/YlQQYUIEAZ76o/giphy.gif)

#CodingIdeal:

```{r,eval=FALSE}
od_test$perc_cycle=round(od_test$bicycle/od_test$all)*100l=o
d_to_sf(od_test,od_data_centroids)
r=stplanr::route(l=l,route_fun=journey)rnet
=overline(r,"bicycle")
```

--

![](https://media.giphy.com/media/3oKIPnAiaMCws8nOsE/giphy.gif)Realit

y

##Transportsoftware- whichdoyouuse?

```{r,echo=FALSE,message=FALSE,warning=FALSE}
u="https://github.com/ITSLeeds/TDS/raw/master/transport-
software.csv"tms=readr::read_csv(u)[1:5]
tms=dplyr::arrange(tms,dplyr::desc(Citations))
knitr::kable(tms,booktabs=TRUE,caption="Sampleoftransportmodellingsoftwareinusebypractitio
ners.Note:citationcountsbasedonsearchesforcompany/developername,theproductnameand'transpor
t'.Datasource:GoogleScholar searches,October2018.",format ="html")
```

##Datascienceandthetidyverse

-InspiredbyIntroductiontodatasciencewithR(availablefree[online](https://r4ds.had.co.nz/))

```{rtds-
cover,echo=FALSE,out.width="30%"}knitr::include_graphics("https://d33wubrfki0l68.cloudfro
nt.net/b88ef926a004b0fce72b2526b0b5c4413666a4cb/24a30/cover.png")
```

##Ageographicperspective

- Seehttps://github.com/ITSLeeds/TDS/blob/master/catalogue.md

- Paperonthe**stplanr**paperfortransportplanning(available[online](https://cran.r-

project.org/web/packages/stplanr/vignettes/stplanr-paper.html))
- IntroductoryandadvancedcontentongeographicdatainR,especiallythe[transportchapte r](https://geocompr.robinlovelace.net/transport.html)(availablefree[online](https://geo compr.robinlovelace.net/))
- PaperonanalysingOSMdatainPythonwithOSMnx(available[onli ne](https://arxiv.org/pdf/1611.01890))

#Gettingsupport

--

Withopensourcesoftware,theworldisyoursupportnetwork!

--

- Recentexample:https://stackoverflow.com/questions/57235601/

--

- [gis.stackexchange.com](https://gis.stackexchange.com/questions)has21,314questions

- [r-sig-geo](https://r-sig-geo.2731867.n2.nabble.com/)has1000sofposts

- RStudio'sDiscoursecommunityhas65,000+postsalready!

--

- Notransportequivalent(e.g.earthscience.stackexchange.comisinbeta)

- PotentialforaDiscourseforumorsimilar:transportisnot(just)GIS

Textcenter_analysis:

```
---
pagetitle:"Analysisbasedontestcentres"author:"
AmolNanaware"
date:"06/12/2019"always_allo
w_html:trueoutput:
  html_document:default
---
```

~~~~
```{rsetup,include=FALSE}knitr::opts_chunk$set(echo
=TRUE)
```

```{r,warning=FALSE,echo=FALSE,include=FALSE}packages
<-c("plotly", "tidyverse")
newPackages<-
packages[!(packages%in%installed.packages()[,"Package"])]if(length(newPackages))install.pac
kages(newPackages)

library(tidyverse)library(plotl
y)

```

```{r,echo=FALSE}load("pas
sfail.RData")passfail<-
passfail%>%
  mutate(totalFails=Fail1+ifelse(is.na(Fail2),0,Fail2),Totalpass=Pass1+ifelse(is.na(Pass2),0,
Pass2))
```



```{r,echo=FALSE}

passfailGroup<-summarise(group_by(passfail,
Centre),Pass1=sum(Pass1),Fail1=sum(Fail1),Total1=sum(Total1),Pass2=sum(Pass2,na.rm=T),
Fail2=sum(Fail2,na.rm=T),Total2=sum(Total2,na.rm=T),Totalpass=sum(Totalpass),totalFails=
sum(totalFails))

passfailGroup<-
mutate(passfailGroup,Pass1prop=Pass1/Total1,Pass2prop=Pass2/Total2,totalPassProp=(Totalp
ass/(Total1+Total2)),totalFailsProp=(totalFails/(Total1+Total2)))

```


```{r,echo=FALSE}
passfailGroup$totalPassProp=round((passfailGroup$totalPassProp*100),digits=2)passfailGrou
p$totalFailsProp=round((passfailGroup$totalFailsProp*100),digits=2)

passFailGroup1<-
passfailGroup[c(1,8)]passFailGroup1$Test <-
"Pass"names(passFailGroup1)<-
c("Centre","Count","Test")passFailGroup2<-
passfailGroup[c(1,9)]passFailGroup2$Test<-"Fail"
names(passFailGroup2)<-
c("Centre","Count","Test")passFailcount<-
rbind(passFailGroup1,passFailGroup2)
```
~~~~

###Analysisbasedontestcentres
<br/>Inthissectionwewillanalysedatafrom2013till2018abouteachtestcentre.Asshowninthe<ahref="https://github.com/NanawareAmol/R-project_Road-safety/blob/master/Result/loc_spread_across_ireland.JPG">map</a>,thetestcentresarespreadacrosstheIrelandandthenumberofcentresismoreinhighlypopulatedareassuchasdublin,corketc.

Thebarchartshowsthetotalnumberofteststhateachcentreperformedandthetotalpassandfailcountsas wellaspercentages.So, basedonthetest counts,thetop3test centreare, *Fonthill(770685)*,*Deansgrade(767484)*,and*Northpoint2(729661)*.Thebotton3centreswhichperformed lesstestsare,*DonegalTown(16315)*,*Cahirciveen(28806)*and *Clifden(38683)*.

~~~
```{r,echo=FALSE,fig.width=9,fig.height=4}
p<-
plot_ly(passfailGroup,x=~passfailGroup$Centre,y=~passfailGroup$Totalpass,type='bar',name='
Pass',
text=paste("Totaltests=",(passfailGroup$Totalpass+passfailGroup$totalFails),"<br>Passed=",pa
ssfailGroup$totalPassProp,"%","<br>Failed=",passfailGroup$totalFailsProp,"%"),opacity=0.5,m
arker=list(color= '#3AC3E3',line= list(color='#0D6EB0',width=
1)))%>%add_trace(y=~passfailGroup$totalFails,name ='Fails',opacity=0.5,
          marker=list(color='#0E84FF',line=list(color='#0D6EB0',width=1)))%>%layout(yaxis=li
  st(title='Count'),xaxis=list(title='TestCentres'),barmode='stack')
p
```
~~~

####<b>Totaltestpassed foreachtestcentre</b>
Thefollowingscatterplotshowthetotaltestpasscountforeachtestcentrefromtheyear2013tillyear2018.Thequestionsthatcanbeansweredbythisgraphare,<br/>
1. whicharethetop3andlast3centresbasedontotalpasscount?<br/>
&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;<b>(Deansgrade,Northpoint2,FonthillandCahirciveen,Clifden,derrybegresp.)</b><br/>
2. Whichyearhasthehighestandlowesttotalpass count?<br/>
&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;<b>2015and2014respectively</b><br/>But,inthisgraphwearenotconsideringthetotaltestsperformedbythetestcentreswhichshowstheactualperformance ofthetests.Forthiswewillplotanothergraph.
<br/><br/>

~~~
```{r,echo=FALSE,fig.width=9,fig.height=4}
#scatterplotforcentretotalpassperyear
ggplot(data=passfail,aes(x=fct_reorder(Centre, -Totalpass),y=Totalpass,color=Year,size
=Totalpass))+geom_point(alpha=0.5)+
  theme(axis.text.x=element_text(size=9,angle=-
90,hjust=0,vjust=0.5),axis.ticks.x=element_blank(),panel.background=element_rect(fill="
white",colour="lightblue"),panel.grid.minor=element_line(size=0.5,linetype='solid',colour
="lightblue"))+labs(x="TestCentres",y="Totolpasscount")
```
~~~

~~~
####<b>Testperformanceforeachtestcentre</b>
Thegraphgivestheoverallideaofthetestperformancebasedonpassrateandtheyear.
Asper thegraph we cansay thatfor year2013, 2015,2016,2017 and2018,thepassrate
ishigherthat55%.Andthehighestandlowestperformancefound
inKilkennyandMonaghantestcentresrespectively.<br/><br/>
```{r,echo=FALSE,fig.width=9,fig.height=4}
passfail$totPassPercentage<-
round((passfail$Totalpass/(passfail$Totalpass+passfail$totalFails))*100,digits=2)

passfail$totFailPercentage<-
round((passfail$totalFails/(passfail$Totalpass+passfail$totalFails))*100,digits=2)

#scatterplotforcentrepasspercetageperyear
ggplot(data=passfail,aes(x=fct_reorder(Centre,-totPassPercentage), y=totPassPercentage,color
=Year,size=totPassPercentage)) +geom_point(alpha= 0.5)+
  theme(axis.text.x=element_text(size=9,angle=-
90,hjust=0,vjust=0.5),axis.ticks.x=element_blank(),panel.background=element_rect(fill="
white",colour="lightblue"),panel.grid.minor=element_line(size=0.5,linetype='solid',colour
="lightblue"))+labs(x="TestCentres",y="TotalPass%")#title="Test centrepass%peryear",
```

####<b>Totalpasscountlimitsperyear</b>
Theboxplotshowsthetotalpasscountagainsteachyear.Withthiswecan fetchthedetailsonmaximum
andminimumpasscounts per year, themeadian passcountand the
oustandingpasscountvalueswhichareshownasoutliers(points)peryearwiththetestcentrename.<br
>
```{r,echo=FALSE,fig.width=9,fig.height=4}
p<-
plot_ly(passfail,x=passfail$Year,y=passfail$Totalpass,color=~passfail$Year,type="box",text=pa
ste("Centre=",passfail$Centre))%>%
  layout(title="Yearlyperformance",yaxis=list(title='TotalPassCount'),xaxis=list(title='Year'))
p
```
~~~

Task_Rmd :

~~~
---
title:"Project–SignswithSmart ConnectivityforBetterRoadSafety"

output:html_document
---

```{rsetup,include=FALSE}knitr::opts_chunk$set(echo
=TRUE)
#ggparcoord
#geom_polygon=>packcircles

```
~~~

~~~
##Libraries

```{rcars}suppressMessages(library(readxl))
suppressMessages(library(dplyr))suppressM
essages(library(tidyr))suppressMessages(lib
rary(ggplot2))suppressMessages(library(M
ASS))suppressMessages(library(GGally))su
ppressMessages(library(ggExtra))suppress
Messages(library(plotly))suppressMessages(
library(packcircles))
~~~
##PreparingTheDataForAnalysis
```{r}
df<-read_excel("mmAll.xlsx")
#d13<-read_excel("m_m2013.xlsx")
#d14<-read_excel("m_m2014.xlsx")
#d15<-read_excel("m_m2015.xlsx")
#d16<-read.csv("m_m2016.csv",header =T)
#d17<-read_excel("m_m2017.xlsx")
#d18<-read_excel("m_m2018.xlsx")


names(df)[9]<-
"VehicleandSafetyEquipment"names(df)[10] <-
"VehicleandSafetyEquipment%"names(df)[22]<-
"ChassisandBody%"names(df)[26]<-
"SuspensionTest%"names(df)[36]<-
"IncompleteTests%"

df$reportYear<-as.factor(df$reportYear)
```
##Whichpartfailedthemostperreport year?
```{r,echo=FALSE}
######DATA#######
cols<-
c("VehicleMake","VehicleandSafetyEquipment","LightingandElectrical","SteeringandSuspensio
n","BrakingEquipment","WheelsandTyres","Engine,NoiseandExhaust","ChassisandBody",
"SideSlipTest","SuspensionTest","Lighttest","BrakeTest","Emmissions","OTHER")

m<-
df%>%dplyr::select(c("reportYear",cols))%>%group_by(reportYear)%>%summarise_if(is.num
eric,mean,na.rm=TRUE)
m<-gather(m,-reportYear,key=Part,value=Failures)

```
```
```

~~~
```{r}
######PLOT#######
#ggplot(m,aes(x=factor(reportYear),y=,colour=supp,group=supp))+geom_line()

library(MASS)library(GGally
)
#Vectorcolorlibrary(RColorB
rewer)
palette<-brewer.pal(3,"Set1")
 my_colors<-palette[as.numeric(m$reportYear)]

#names(x)<-c("2013","2014","2015","2016","2017","2018")
#p<-ggparcoord(m,
columns=2:13,groupColumn="reportYear")+geom_line(size=0.3)+theme_minimal()+geom_point(
)+
#xlab("CarPart")+ylab("Averagefailurerate")


ggplotly(ggplot(data=m,
  mapping=aes(x=reportYear,y=Failures,colour=Part,group=1))+geom_point()+
    geom_line()+xlab("ReportYear")+ylab("AverageNumberofFailures")

)
```
##EquipmentFailures-OverallStatistics

```{r}library(ggplot2
)
cols<-
c("VehicleandSafetyEquipment","LightingandElectrical","SteeringandSuspension","BrakingEqui
pment","WheelsandTyres","Engine,NoiseandExhaust","ChassisandBody","SideSlipTest","Suspen
sionTest","Lighttest","BrakeTest","Emmissions","OTHER")
a<-
df%>%dplyr::select(cols)b<-
colSums(a)
c<-data.frame(Part=names(b),Percent=unname(b)/sum(df$Total)*100)ggplot(c)+
  geom_col(mapping=aes(x=reorder(Part,-
Percent),y=Percent,fill=Percent),col="black")+xlab("")+
  ylab("FailurePercentage(%)")+scale_fill_gradient(low
  ="orange",high="tan")+coord_flip()

```
###Thereisabuginthis code.Cananybodyfix it?
```{r,eval=FALSE}
####Thepolygongraphrepresentationoftheabovedata####l<-
data.frame(Part=names(b),Total=unname(b))
packing<-circleProgressiveLayout(l$Total,sizetype='area')
~~~

~~~
l$packing<-packingl
dat.gg<-circleLayoutVertices(packing,npoints=50)

p<-ggplot()+geom_polygon(data=dat.gg,aes(x,y,group=id,fill=as.factor(id)),colour="black",
alpha=0.6) +geom_text(data=l, aes(x, y,
size=Total,label=Part))+scale_size_continuous(range=c(1,4))+theme_void()+theme(legend.posit
ion="none")+coord_equal()



ggplotly(p,tooltip=c("Total","Part"))

```
```{r}
z<-
df%>%group_by(VehicleMake)%>%summarise(tot=sum(Total),res=sum(PASS)/sum(Total))%
>%arrange(desc(tot))%>%print(Inf())
```

```
{r}require(scales
)
q<-z%>%arrange(desc(tot))%>%slice(1:15)

ggplot(q)+
  geom_col(mapping=aes(x=reorder(VehicleMake,-
tot),y=tot,fill="green"))+xlab("VehicleMake")+ylab("NumberofVehicles")+coord_flip()+theme(
legend.position="none")+
  scale_y_continuous(labels=comma)

#ggMarginal(g,type="histogram",fill="transparent")
```

##PassPercentageversusNumberofVehiclesforagivenVehicleMake
```{r}require(scal
es)library(plotly)
p<-ggplot(q,aes(x=tot,
  y=res*100))+geom_line(color="red")+
  geom_point(aes(text=VehicleMake))+xlab("NumberofVehicles")+ylab("PassPercentage(%)")+
  scale_x_continuous(labels=comma)

ggplotly(p,tooltip="text")
```
~~~