Prash_survey_analysis

prash

2018-04-09

```
library(tidyverse)
library(knitr)
library(stringr)
```

```
survey <- read_csv("../data/survey_responses_latest.csv")
```

```
kable(head(survey))
```

Timestamp	This online survey company is hosted by a web survey company located in the USA and as
2018/04/09 2:30:02 PM MDT	I agree
2018/04/09 2:30:05 PM MDT	I agree
2018/04/09 2:45:19 PM MDT	I agree
2018/04/09 2:53:19 PM MDT	I agree
2018/04/09 3:05:15 PM MDT	I agree
2018/04/09 3:05:50 PM MDT	I agree

```
survey_analysis <- survey %>%
  select(-c(1,2))


colnames(survey_analysis) <- c("academic","codingexp","lovecoding","firstprog","preferprog","dstask","n
```

```
kable(head(survey_analysis))
```

academic	codingexp	lovecoding	firstprog	preferprog	dstask	noofprog
Business	Less than 1	Yes	SAS	Python	Data visualization	3
Electrical Engineering	Less than 1	Yes	C	Python	Machine Learning	3
Geography	Less than 1	Indifferent	R	R	Data visualization	2
Mathematics / Statistics	1 to 5	Indifferent	Java	Python	Machine Learning	2
Mathematics / Statistics	Less than 1	Indifferent	Pascal	R	Data wrangling	2
Engineering	1 to 5	Yes	Java	R	Data visualization	2

```
survey_analysis %>% group_by(preferprog)%>%
  summarize(count=n()) %>%
  kable()
```

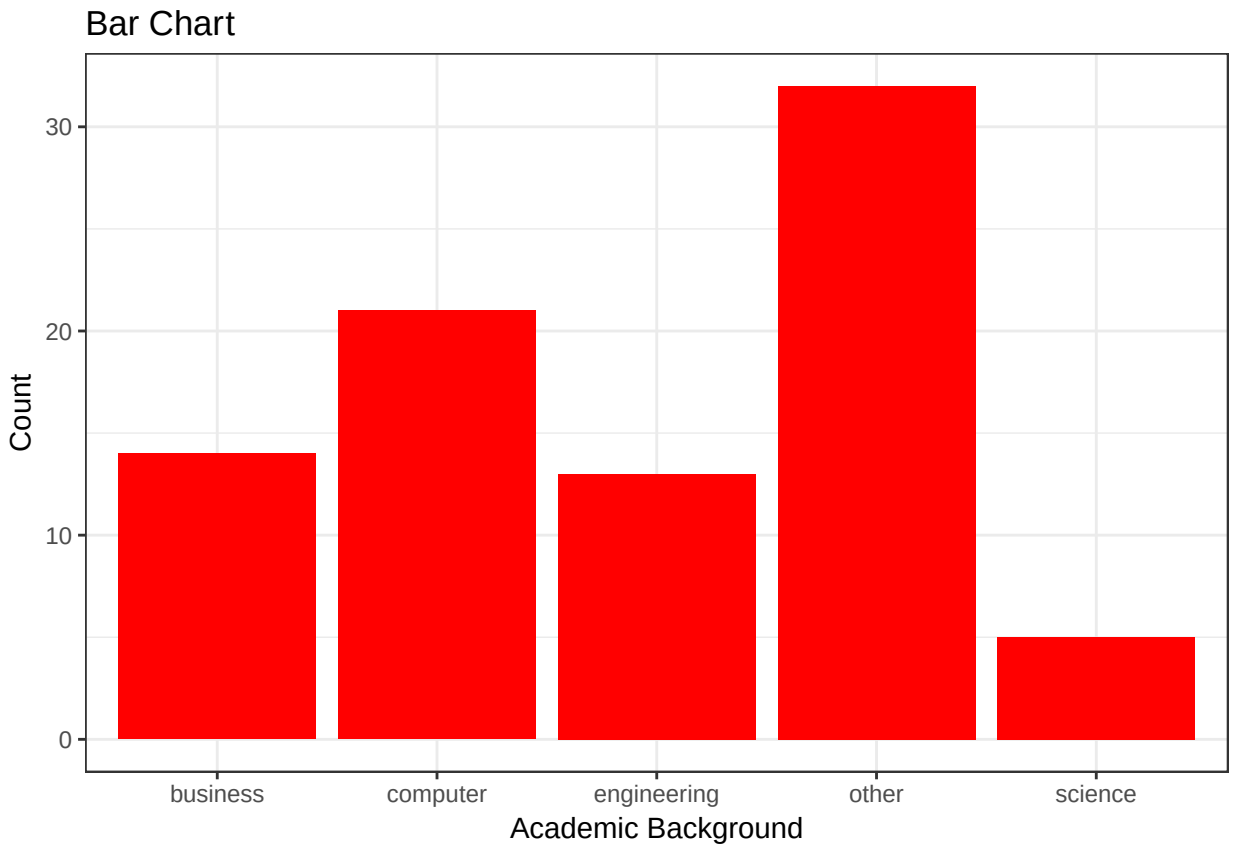
preferprog	count
Python	55
R	30

```
#survey_analysis$academic <- str_replace(tolower(survey_analysis$academic),c(".*econ.*",".*fin.*",".*bu
```

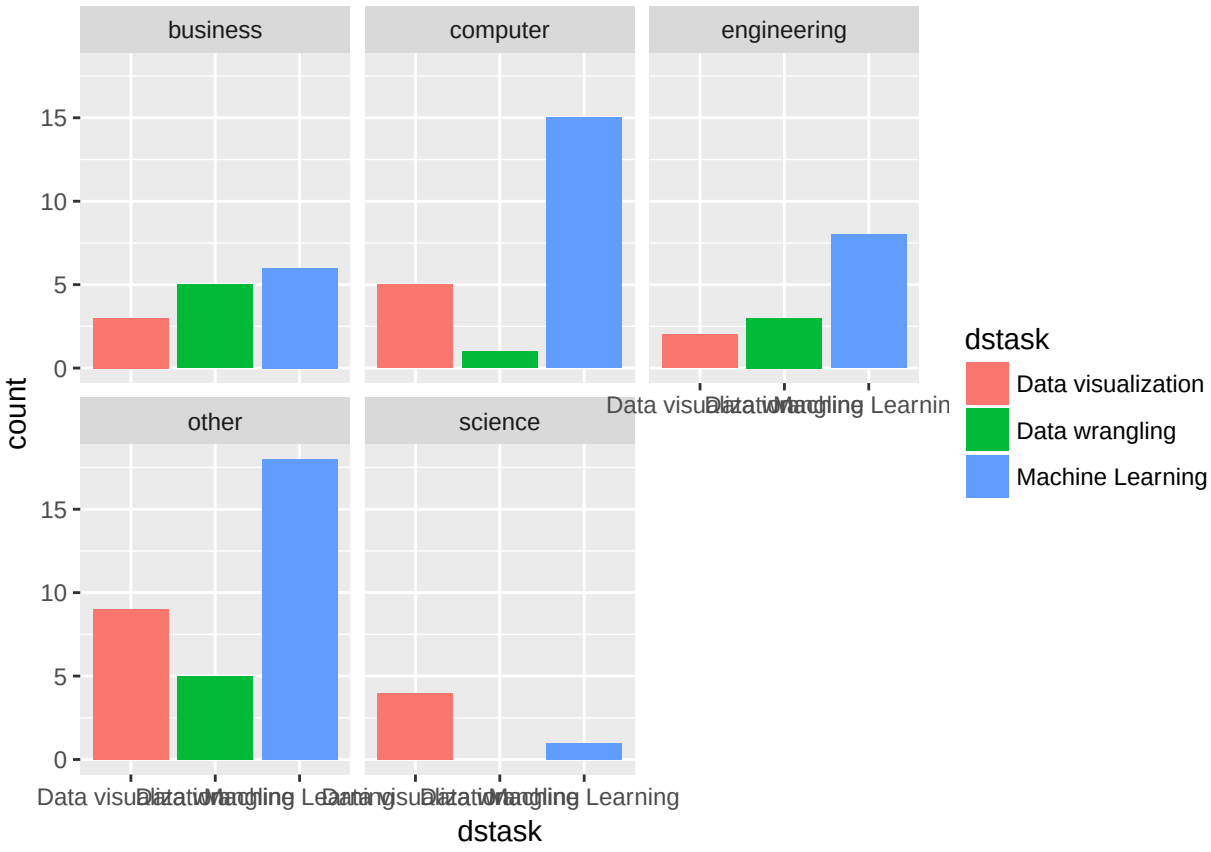
```r
#survey_analysis$academic <- str_replace(tolower(survey_analysis$academic),c(".*econ.*",".*fin.*",".*bu

#?str_replace

survey_analysis$academic <- str_replace(survey_analysis$academic,"Computer Science / Computer Engineeri

survey_analysis$academic <- str_replace(tolower(survey_analysis$academic),".*econ.*","business")

survey_analysis$academic <- str_replace(tolower(survey_analysis$academic),".*busi.*","business")

survey_analysis$academic <- str_replace(tolower(survey_analysis$academic),".*eng.*","engineering")

survey_analysis$academic <- str_replace(tolower(survey_analysis$academic),".*science.*","science")

# Creating a vector, putting all the fields to others

survey_clean_1 <- c("business","computer","engineering","science","Mathematics / Statistics")

#interact_clean$name[!(interact_clean$name %in% interact_clean_1)] <- "other"

survey_analysis$academic[!(survey_analysis$academic%in% survey_clean_1)] <- "other"


# Summary
survey_analysis_count<- survey_analysis %>%
  group_by(academic) %>%
  summarise(count=n())

kable(survey_analysis_count)
```

academic	count
business	14
computer	21
engineering	13
other	32
science	5

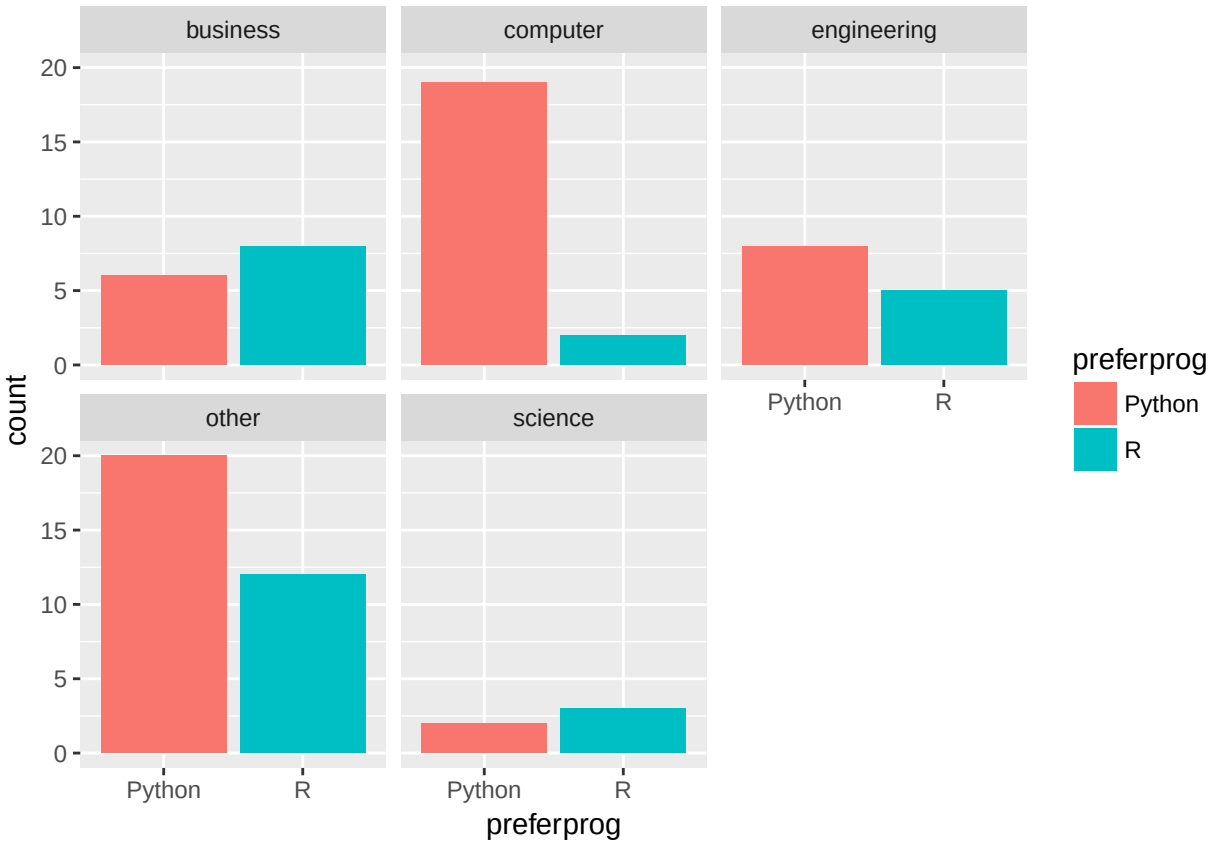
```r
survey_analysis %>%
  ggplot(aes(academic))+geom_bar(fill="red",position = "dodge")+
  labs(title="Bar Chart",x="Academic Background",y="Count") +
  theme_bw()
```




```
survey_analysis %>%
 ggplot() +
 geom_bar(aes(x = dstask,fill = dstask)) +
 facet_wrap(~academic)
```




```
survey_analysis %>%
 ggplot() +
 geom_bar(aes(x = preferprog,fill = preferprog)) +
 facet_wrap(~academic)
```


business
computer
engineering
other
science
count
20
15
10
5
0
Python
R
preferprog
preferprog
Python
R