IIT Admission Trend (2016-2022

Kamleshwar

2023-06-12

Project

I analyse admission trend, not only of India's but worlds best insitute IIT from 2016 to 2022

Reading data

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(plotrix)
setwd("C:/Users/Admin/Downloads/archive/")
csv<- read.csv("student_data.csv",header = TRUE, sep = ",")
csv_tbl<- as_tibble(csv)
head(csv_tbl)
```

```
## # A tibble: 6 × 10
##    Stude…¹ Stude…² Date.…³ Field…⁴ Year.…⁵ Expec…⁶ Curre…⁷ Speci…⁸   Fees Disco…⁹
##      <int> <chr>   <chr>   <chr>     <int>   <int>   <int> <chr>    <int>   <int>
## 1  165527 Bryan … 2006-0… Comput…    2020    2017       3 Web De… 155152   19572
## 2  635763 James … 1999-0… Mechan…    2020    2020       2 Machin… 157870   14760
## 3  740021 David … 1997-1… Civil …    2017    2022       1 Networ…  55662    5871
## 4  433076 Susan … 1999-1… Comput…    2021    2019       1 Data S… 134955   17284
## 5  441628 Britta… 1998-0… Chemic…    2016    2018       1 Networ… 125934   14871
## 6  837685 Melody… 1997-1… Mechan…    2016    2019       4 Web De… 131633    7565
## # … with abbreviated variable names ¹Student.ID, ²Student.Name, ³Date.of.Birth,
## #   ⁴Field.of.Study, ⁵Year.of.Admission, ⁶Expected.Year.of.Graduation,
## #   ⁷Current.Semester, ⁸Specialization, ⁹Discount.on.Fees
```

```
#to see all column heading
colnames(csv)
```

```
##  [1] "Student.ID"          "Student.Name"
##  [3] "Date.of.Birth"       "Field.of.Study"
##  [5] "Year.of.Admission"   "Expected.Year.of.Graduation"
##  [7] "Current.Semester"    "Specialization"
##  [9] "Fees"                "Discount.on.Fees"
```

Creating new column

we create new column that we use for analysis

```
#adding new column in dataset
discount_percentage<- (csv$Discount.on.Fees/csv$Fees)*100
discount_percentage<- round(discount_percentage,2)
csv$discount_percentage<-(discount_percentage)
#again checking all column heading
colnames(csv)
```

```
##  [1] "Student.ID"          "Student.Name"
##  [3] "Date.of.Birth"       "Field.of.Study"
##  [5] "Year.of.Admission"   "Expected.Year.of.Graduation"
##  [7] "Current.Semester"    "Specialization"
##  [9] "Fees"                "Discount.on.Fees"
## [11] "discount_percentage"
```
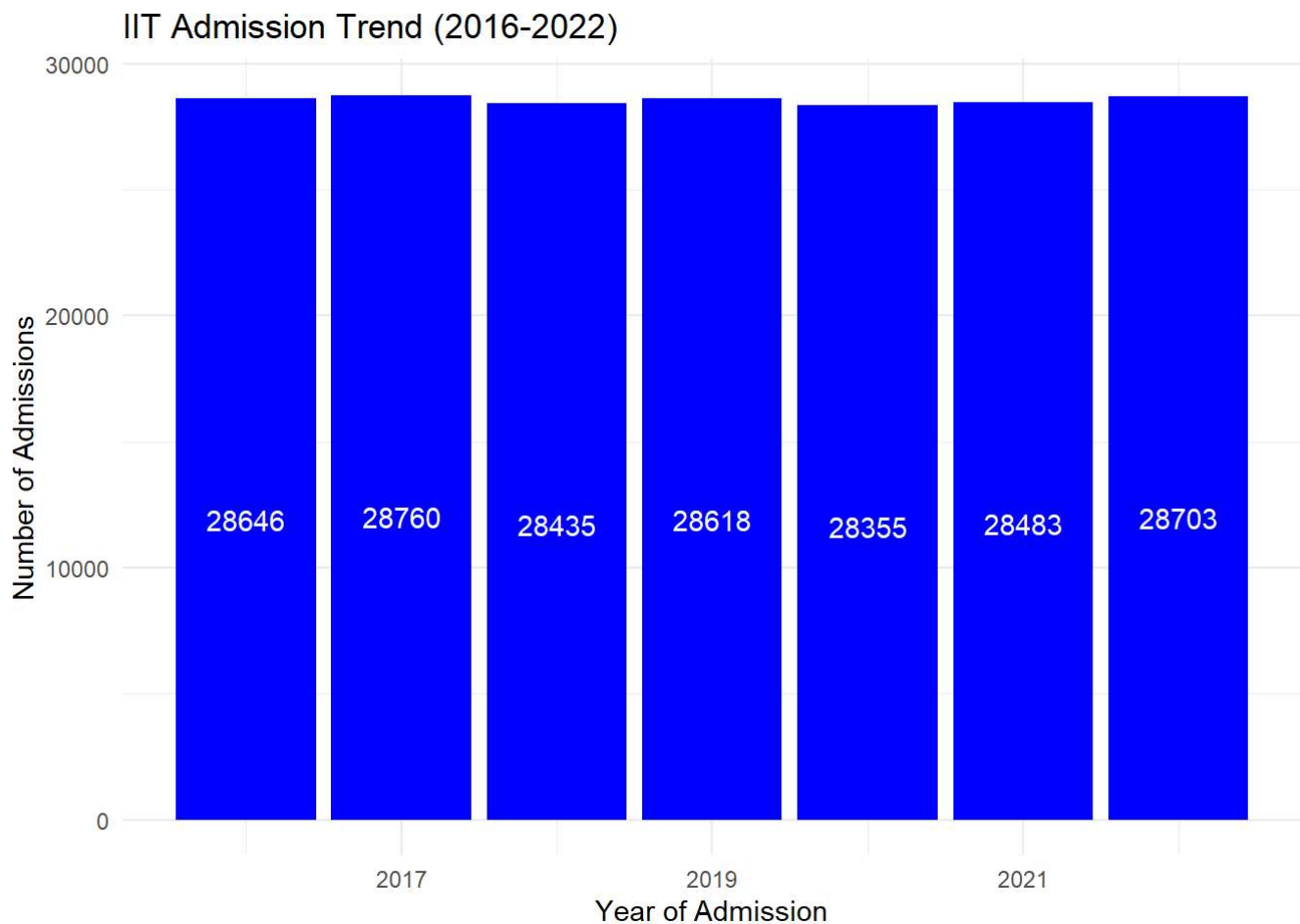
Admision per year

Now we check admission taken per year in IIT

```
admission_per_year<-(count(group_by(csv,Year.of.Admission)))
print(admission_per_year)
```

```
## # A tibble: 7 × 2
## # Groups:   Year.of.Admission [7]
##   Year.of.Admission     n
##             <int> <int>
## 1              2016 28646
## 2              2017 28760
## 3              2018 28435
## 4              2019 28618
## 5              2020 28355
## 6              2021 28483
## 7              2022 28703
```

now we plot the graph

Including Plots

IIT Admission Trend (2016-2022)

as we can see admission taken per year is almost around at 28,000

Poular Department

Now we check which department have most student

```
field_with_most_students<-(count(group_by(csv,Field.of.Study)))
print(field_with_most_students)
```

```
## # A tibble: 5 × 2
## # Groups:   Field.of.Study [5]
##   Field.of.Study           n
##   <chr>                <int>
## 1 Chemical Engineering  40020
## 2 Civil Engineering     39774
## 3 Computer Science      39960
## 4 Electrical Engineering 40305
## 5 Mechanical Engineering 39941
```

lets plot graph

## IIT Admission Trend (2016-2022)
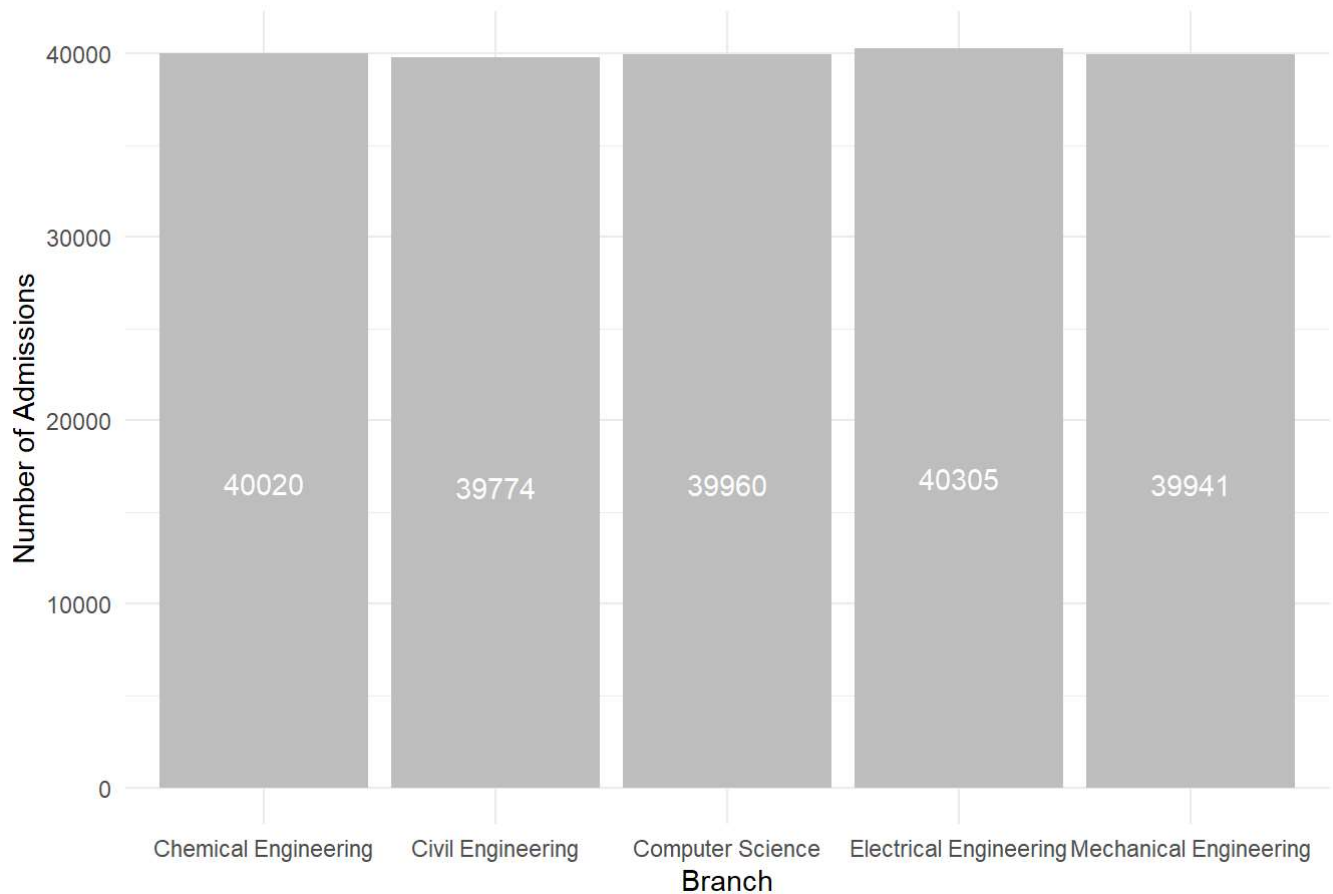


now we check pie chart but firt we create a data frame that calculate percentage

```
department_percentage= (field_with_most_students$n/sum(field_with_most_students$n))*100
department_percentage=round(department_percentage,2)
department_percentage<-data.frame(Department = field_with_most_students$Field.of.Study, Percenta
ge = department_percentage)
department_percentage
```

```
##                 Department Percentage
## 1    Chemical Engineering      20.01
## 2       Civil Engineering      19.89
## 3        Computer Science      19.98
## 4 Electrical Engineering      20.15
## 5 Mechanical Engineering      19.97
```

now let see pie chart

19.89

20.01

19.98

19.97

20.15

	Chemical Engineering
	Civil Engineering
	Computer Science
	Electrical Engineering
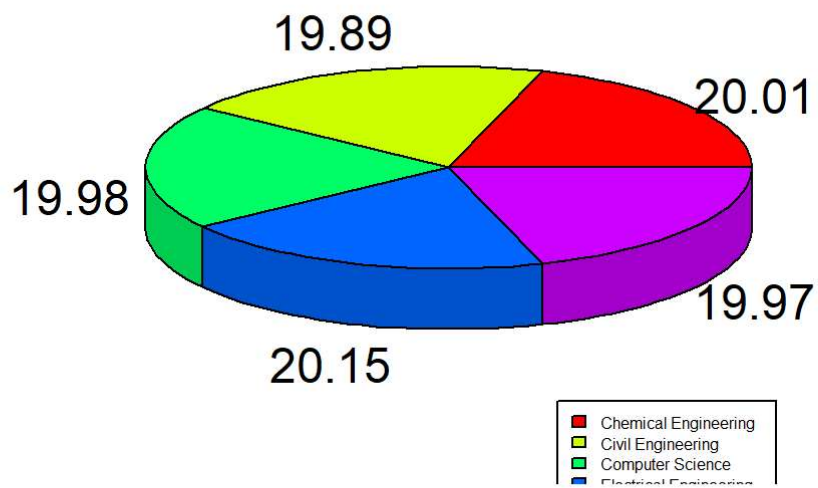
Comparing Fess and Discount
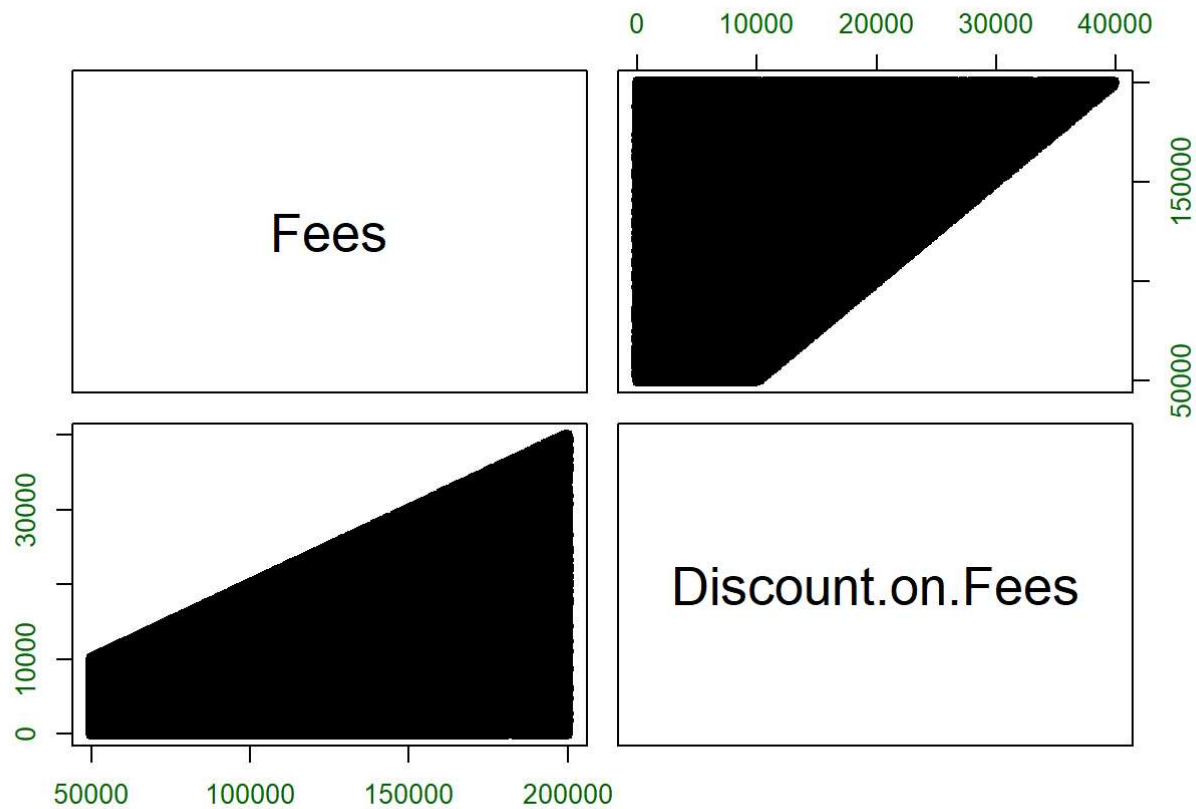
now we check maximum and minimun fees and discount respectivly but first we have to create a dat frame

```
data <- data.frame(
  category = c("Fees", "Discount"),
  value = c(max(csv$Fees), max(csv$Discount.on.Fees)),
  ymin = c(min(csv$Fees), min(csv$Discount.on.Fees)),
  ymax = c(max(csv$Fees), max(csv$Discount.on.Fees))
)
```

Maximum and Minimum Fees and Discounts