An investigation on travel pattern of households of Germany using software R

Group-6

Abu-Er-Rubb Rami Jordanien

Antonio Yarid Mexiko Bedolla Hernandez

Mahbub Alam

Martins Tamovics

Md. Reza Qamran Choudhoury

Zain-ul-Abedin

Prof. Regine Gerike, mobil.TUM

Abstract

The following paper analyses some aspects of the mobility in Germany based on data collected from the transport research part of the "Mobilität in Deutschland 2008 – MID 2008" study. The population sample of 50,000 households helped in getting results that could be adopted for the region. Statistical tests will be used to answer some research questions that drew our attention. The first question is whether there is a strong relation between the economical status of households and the number of cars they own. The second question is about the difference, between the population of the eastern and western states of Germany, in the attitudes towards private vehicles. T-test will be done to check if the difference between the mean number of cars per person, in the east and west, for the same income levels is significant. The test shows a mix attitude. The third question is about which group of people, based on gender, age, education level, employment and activity is the most dependent on public transport. It has been observed that level of income influence car ownership and mean ownership of car in west and east part for same income level is quite same.

Table of Contents

1. Introduction

Mobility is a vital part of our daily life. It's the back bone of modern economies and societies; no one can imagine current global village as immobile. In this paper, the data set is quite huge and there are too many variables, thus giving us the ability to analyze the travelling patterns which depends on a number of different factors like education, economical conditions, number of people in a household, which part of the country people are living. There are many more factors but we select some of the main variables in the data set to answer the general questions, whom answers are taken for granted in the society.

As the general perception about rich people is that they will buy many and expensive cars, they live a very fancy life, they are the least, who uses the public transport, but many others factors also influence these behaviors as well. For instance, people living in far flung areas not in the city rather outside of the city, there is not enough public transport available or Government can't build roads, rails tracks for small number of people, due to many limitations from politics to bureaucracy and from technology to resources available. But number of cars per house hold depending upon the economic status can explain this general question about rich people own more cars than power people.

East part of the Germany as compared to its west part is less developed and people usually have less monthly incomes, higher unemployment rates, may be less no of cars ownership in east part then the west part. In the contrary the west part is more developed, more job opportunities, good infrastructure, good public transport system and more economic prosperity in this part. These factors can play a vital role in deciding the number of cars per household have.

Different group segments uses different mode of transport, For instance mostly students use public transport or their own bikes to go to their schools or colleges, People who are doing regular jobs may also use Public transport to get rid of traffic jams to be more productive and

always reach office on time. So by taking some of the factors we can figure out which type of group use public transport most frequently.

2. Methodology

2.1 Car Ownership in Germany

Is there a relationship between the economical status of the household and the number of private vehicles owned?

To understand the mobility pattern of the people of any area, logically we have to investigate the spending psychology of the people. People may spend more money to have a comfortable ride without considering the big picture or can spend less and satisfied with the available services. Also, we have to understand the allocation of their income to various household needs. In which part do they spend more or for our sake of interest we can ask, what portion of their income they spend on travel. Simply, we can get a good picture by finding a relation between income of households and number of vehicles owned by corresponding households.

It is a very simple relation but to understand the mobility pattern, we realize its importance. No doubt one of the main parameters for traffic analysis is the number of trips made with a private vehicle. While it is usually the highest share of vehicles running on roads, the car has the lowest passenger capacity/energy consumption ratio. People tend to favor this mode of transport due to the time flexibility it gives, the level of comfort, the reduction of walking distance, the abolishment of waiting time at stations, and the time they get to spend on their own in it listening to their favorite music or simply sitting in silence and relaxing. Of course, this is only a valid option when the people can afford owning a private vehicle.

Now we can return to the question whether everyone owns a private vehicle for the previously mentioned reasons which are somehow justified or some own them for other reasons like showing off their economic status or for the excitement they get when driving powerful cars.

As we have been supplied with survey data for almost 26,000 households, we have access to the economical status (income) of each household along with the number of vehicles owns. It is a good idea to use income and number of vehicles per individual rather than households to discuss in uniform unit as the number of persons per household is not same. But there arise a

different issue. We do not know the distribution of income among the family members. In order to simplify the matter, we can ignore the size of the households and assume it is uniform.

We can categorize these households in 15 income ranges and find number of vehicles within these ranges. It has been observed in fig. (1) That, with increase in income, there is increase in number of vehicles per households. But after income level <2600 Euros, we notice that people with income level greater than 2600 Euros/month have 2 or more cars and less households having 1 car. Moreover, the most frequent monthly income groups (100, 2000 & 2600) having a single car per household. Also with income groups greater than 5000 Euros/month have mostly 2 cars or more and having a single car is the least.

2.1.1 Method:

To performer this Procedure, we need to separate the data into different parts depending on the household income level and number of cars per household. We have 15 different level of income starting from Euro 500 to 7000 or more. First we remove all the false (N.A) values which are not required for our procedure.

Next step was to count the number of cars for particular household's income level in whole Germany dataset. The amount of counted cars and households income levels were written into the data variables for the further use for t test and Graphs.

Afterwards we use these data variables to plot main graphs and sub graphs per income level. Finally we do some T-Tests. Further explanation has been provided in R-code in section 6.1

2.1.2 Descriptive Analysis:

Main Graph

The main graph is plotted through using h2008$hheink and h2008$h04_3 variables, this graphs shows number of cars per income levels and income levels are divided in to 15 groups (500-7500).

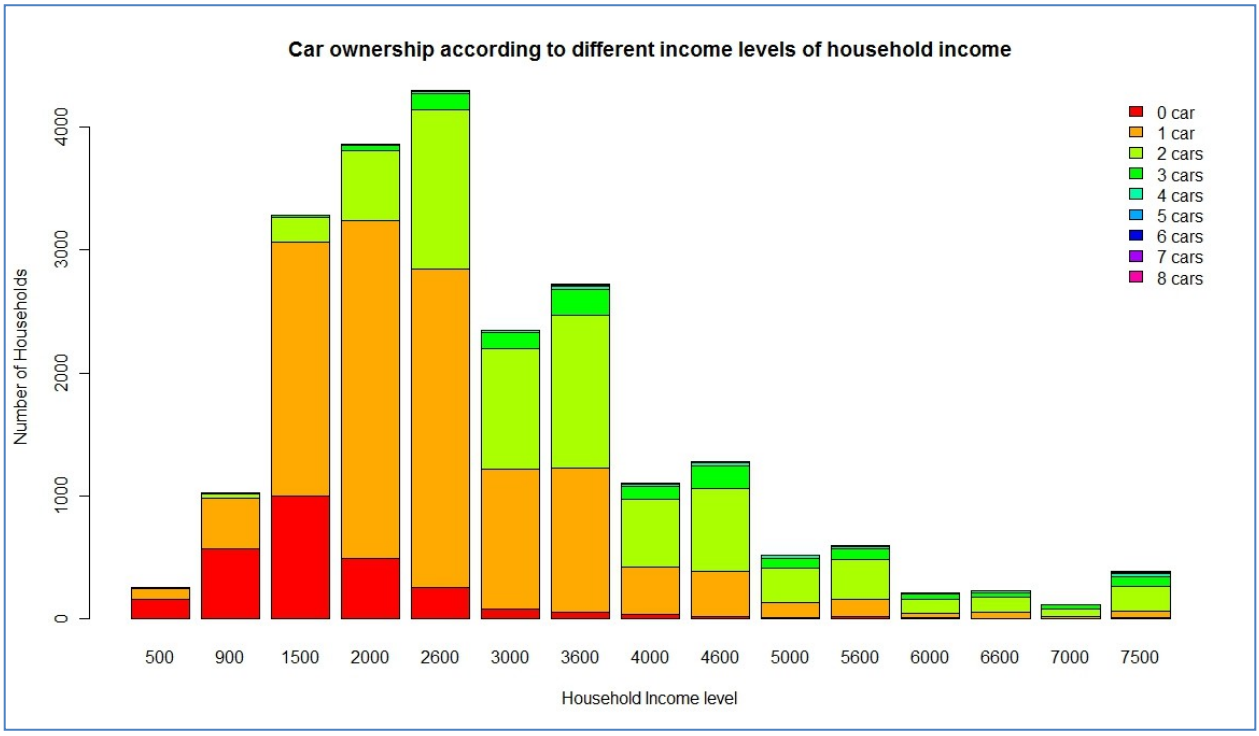


Figure 1: Number of cars per household depending upon the income level

Number of cars Per Income level

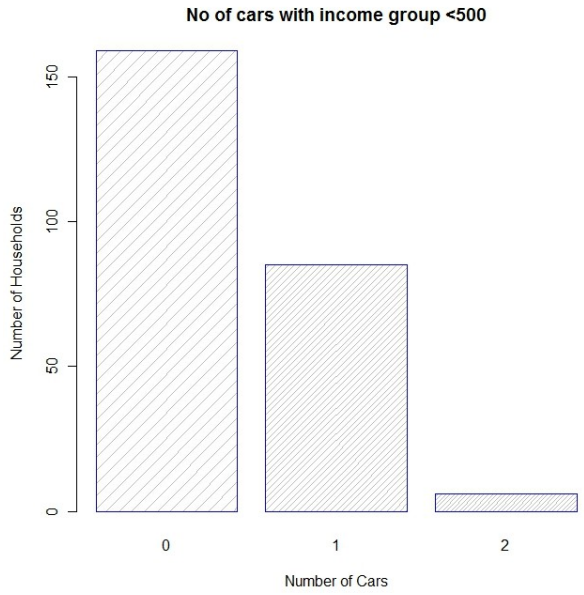


Fig 2:Number of cars in income group level <500 Euros

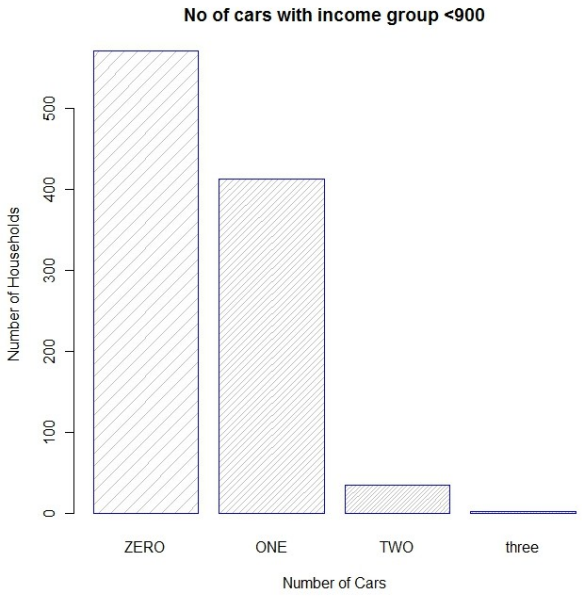


Fig 3:Number of cars in income group level <900 Euros

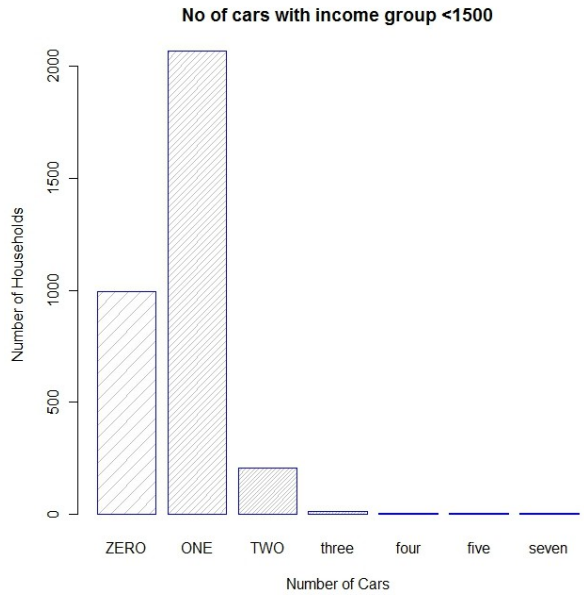


Fig 4:Number of cars in income group level <1500 Euros

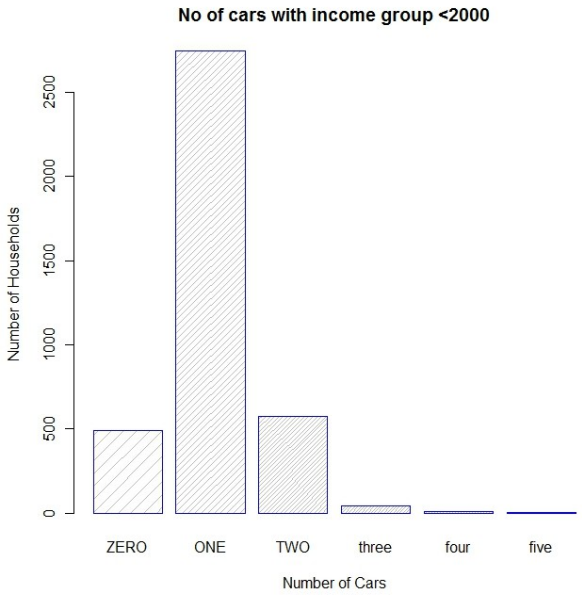


Fig 5:Number of cars in income group level <2000 Euros

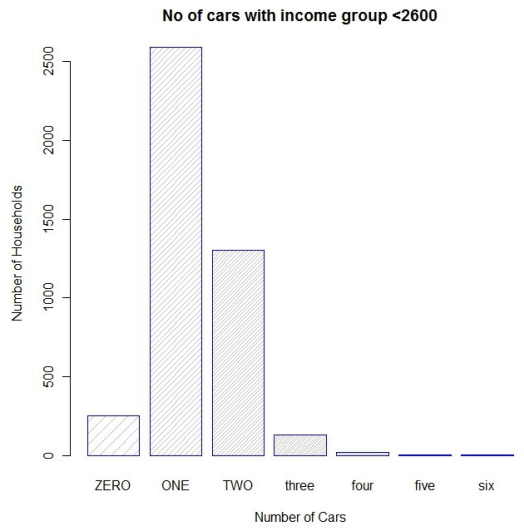


Fig 6:Number of cars in income group level <2600 Euros

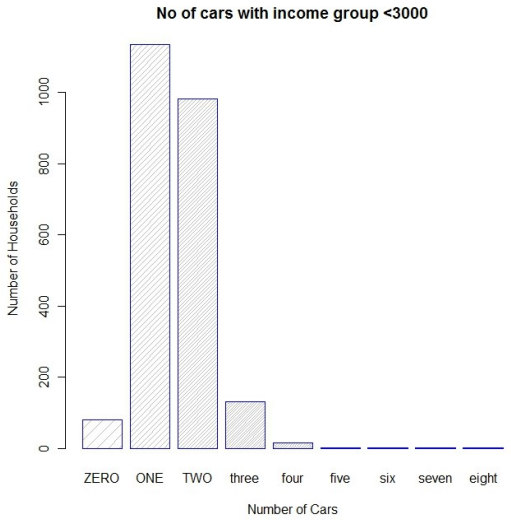


Fig7:Number of cars in income group level <3000 Euros

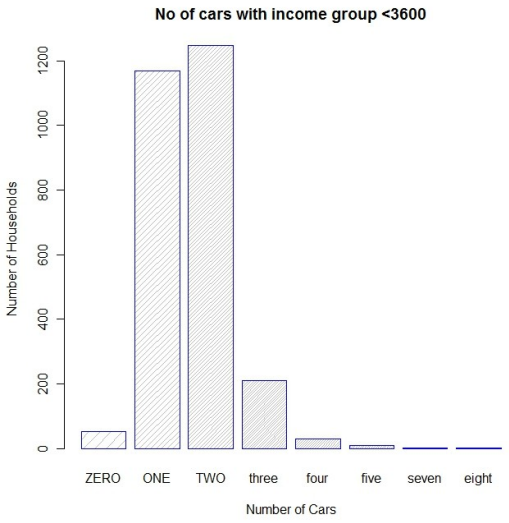


Fig 8:Number of cars in income group level <3600 Euros

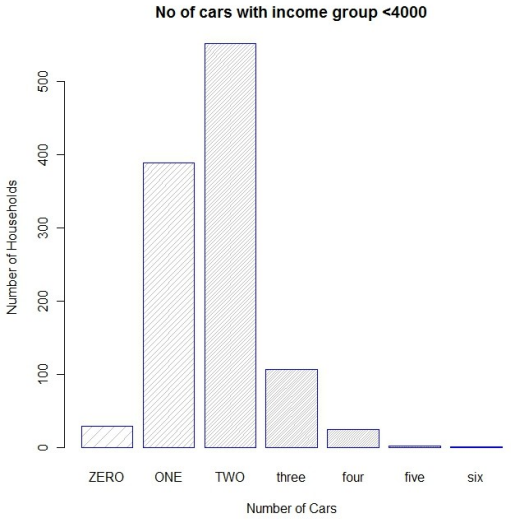


Fig 9:Number of cars in income group level <4000 Euros

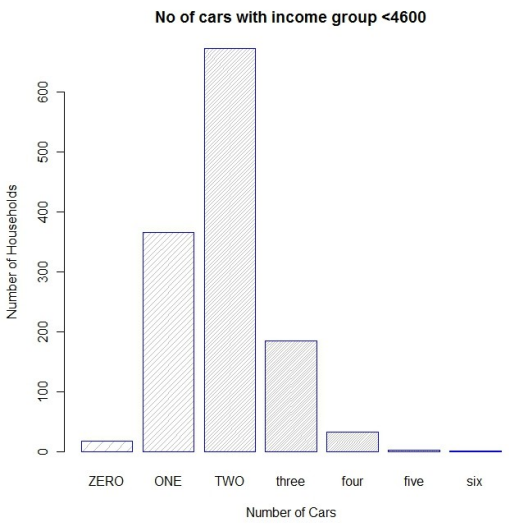


Fig 10:Number of cars in income group level <4600 Euros

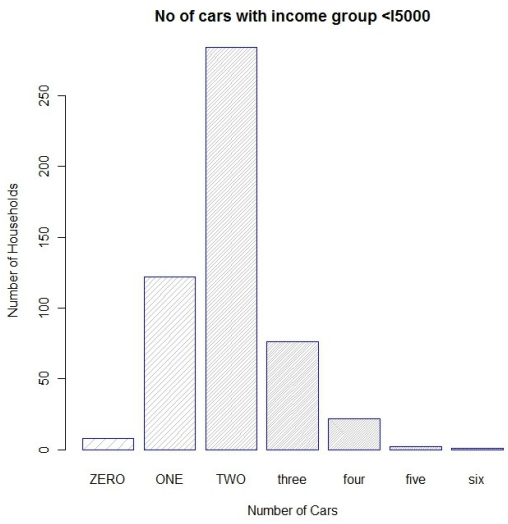


Fig 11:Number of cars in income group level <5000 Euros

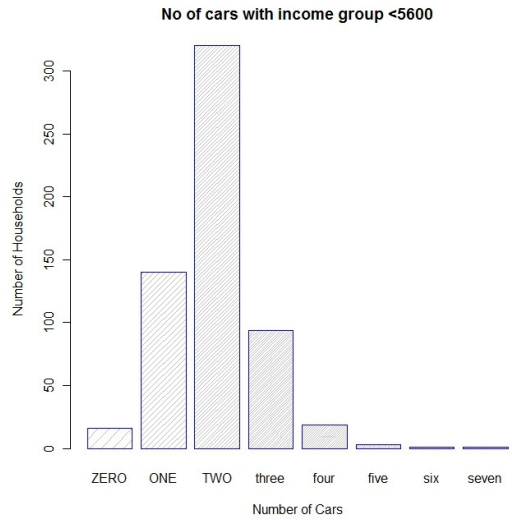


Fig 12:Number of cars in income group level <5600 Euros

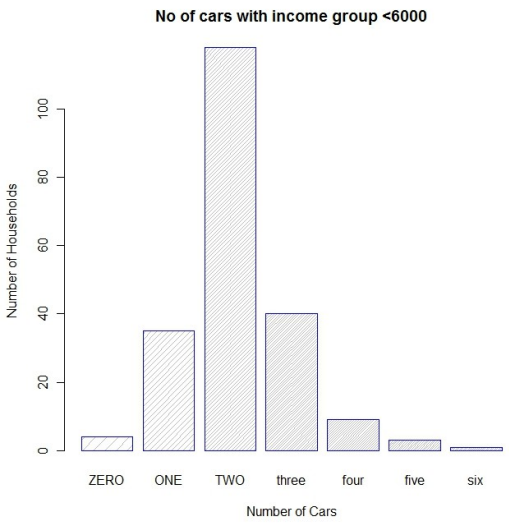


Fig13:Number of cars in income group level <6000 Euros

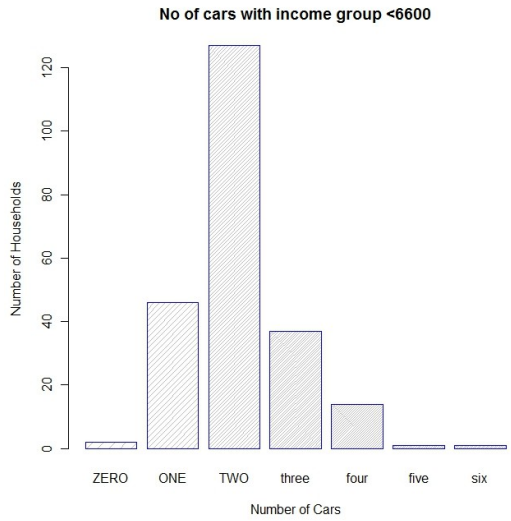


Fig 14:Number of cars in income group level <6600 Euros

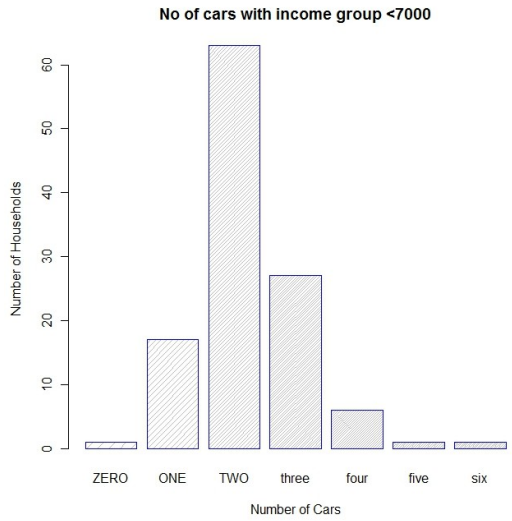


Fig 15:Number of cars in income group level <7000 Euros

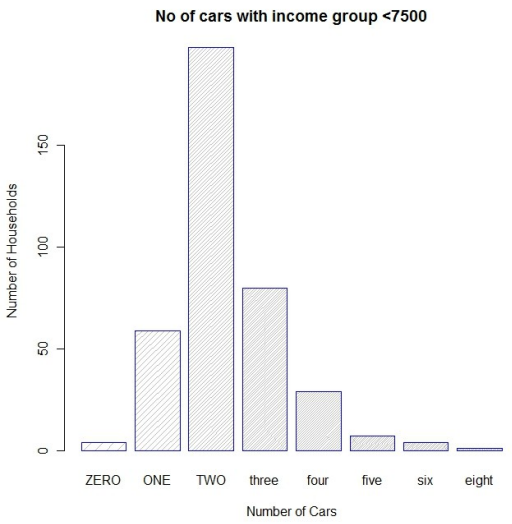


Fig 16:Number of cars in income group level <7500 Euros

2.1.3 T-Tests:

Welch's *t* test is an adaptation of Student's *t*-test intended for use with two samples having possibly unequal variances. As such, it is an approximate solution to the Behrens–Fisher problem.

With the help of T-Tests, one can test the null hypothesis that the two population means are equal (using a two-tailed test), or the null hypothesis that one of the population means is

greater than or equal to the other (using a one-tailed test). In particular, the test will yield a "p-value" which might or might not give evidence sufficient to reject the null hypothesis.

In our case, the tests are all two sided; our objective is to measure the mean car ownership between different income groups (15 income levels). All the tests were performed at a 95 percent confidence interval because it's the default for Project-R. a difference between means only becomes insignificant, ie. the null hypothesis is accepted, if the P-value output is 2.5% or higher. In the following, all income values are per month and in Euro.

1. T-test for an income level <5000 and an income level <7500.

H0: μ1 = μ2 [two sided]	
H0	Null hypothesis
μ1	The mean of the sample 1
μ2	The mean of the sample 2
H0: μ1 ≠ μ2:	The car ownership per household is not the same for an income level <5000 and an income level < 7500.

```
t.test(income10,income15,"two.sided")
Welch Two Sample t-test
data:  income10 and income15
t = -4.9168, df = 714.759, p-value = 1.092e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -0.4393296  -0.1885970
Sample estimates:
mean of x     mean of y
1.984466      2.298429
```

The result shows that the null hypothesis is to be rejected as the difference between the means of the two is significant (1.984, 2.298) at 95% confidence interval and the p-value = 1.092e06.AS from the figure one, both the bars showed a little different pattern of car ownership

2. T-test for an income level < 900 and an income level <4000.

H0: μ1 = μ2 [two sided]	
H0	Null hypothesis
μ1	The mean of the sample 1
μ2	The mean of the sample 2
H0: μ1 ≠ μ2:	The car ownership per household is not the same for an income level <900 and an income level <4000.

```
t.test(income2,income8,"two.sided")
Welch Two Sample t-test data:  income2 and income8
t = -42.7678, df = 2019.543, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: -1.322931  -1.206923
sample estimates:
mean of x     mean of y
0.4789422     1.7438692
```

The result shows that the null hypothesis is to be rejected as the difference between the means of the two is significant (.479, 1.743) at 95% confidence interval and the p-value = 2.2e-16.AS from the figure one, both the bars showed a high difference as more cars per household with income level <4000 then <900.

3. T-test for an income level <2600 and an income level <3600.

H0: μ1 = μ2 [two sided]	
H0	Null hypothesis
μ1	The mean of the sample 1
μ2	The mean of the sample 2
H0: μ1 ≠ μ2:	The car ownership per household is not the same for an income level <2600 and an income level < 3600.

```
> t.test(income5,income7,"two.sided")
data:  income5 and income7
t = -18.3653, df = 5263.382, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:  -0.3581102  -0.2890307
sample estimates:
mean of x    mean of y
 1.323488     1.647059
```

The result shows that the null hypothesis is to be rejected as the difference between the means of the two is significant (1.323, 1.647) at 95% confidence interval and the p-value = 2.2e-16.AS from the figure one, both the bars showed a significant difference as more cars per household with income level <3600 then <2600.

4. T-test for an income level <6000 and an income level <6600.

H0: μ1 = μ2 [two sided]	
H0	Null hypothesis
μ1	The mean of the sample 1
μ2	The mean of the sample 2
H0: μ1 ≠ μ2:	The car ownership per household is not the same for an income level <6000 and an income level <6600.

```
t.test(income12,income13,"two.sided")
data:  income12 and income13
t = 0.4381, df = 430.508, p-value = 0.6615
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:  -0.1284289   0.2021131
sample estimates:
mean of x    mean of y
 2.133333     2.096491
```

The result shows that the null hypothesis is to be accepted as the difference between the means of the two is not significant (2.133, 2.096) at 95% confidence interval and the p-value = 0.6615.AS from the figure one, both the bars showed somewhat similar pattern of car ownership.

2.1.4 Conclusion

•It has been observed that change in income results change in car ownership as well. But we can say that almost 70% households, belonging to any particular income range except below 900€/month and above 3600€/month, posses at least a car.

• For income level above 4000€/month, 50-60% households own at least 2 cars. It is apprehended that one car is for the earning person(s) and the other one serves for the Children or both cars are operated by earning persons only.

• For income range below 900€/month, more than 50% of households don't own any car. Even with income 4000€/month, there are several households (approx. 5%) that don't own a car. It is a different example, maybe the income is not stable or there arise different explanation beyond our interest.

2.2 Mean car ownership investigation in East, West Germany

We may experience different behavior at different places as different regions have different properties e.g. economy, weather conditions, political situations etc. The first question already showed the relationship between the income status of a household and the number of vehicles owned. In this regard, the spatial difference was not considered. But now the difference between those patterns in the East and West of Germany will be checked. It is anticipated that the difference will be useful for future transport planning.

It can be expected, that there is not a big difference between both parts of Germany in general or there can be a difference. However in both cases we will also check the difference between each income group.

The results can either mean that people have different opinions based on their location, the infrastructure and city structure in general is better or worse for implementing public transport or that the advertisement market is not targeting the right groups. The difference of amount of income per household is banned by considering mean number of car ownership.

Getting no significant difference means that there is no need to split the population into two groups when studying other factors.

Households can be separated based on the east and west classification. Having the economic status of each household along with the number of people in it, the average income per household can be calculated. Afterwards, the mean number of cars for all households of a certain economic status in the east will be compared to the mean number of cars for all households of that same economic status in the west. This will be repeated for all income ranges.

2.2.1 Method

To perform this investigation, we need to separate the data into two parts depending on the location in Germany. It was performed by checking the location before checking the income level.

Next step was to count the number of cars and the number of households in each part of the Germany. The amount of counted cars and households were written into the data variable for the further use.

Afterwards necessary result of mean cars per household of each income level in each part of Germany and the mean amount of cars in both sides of Germany were calculated. Further explanation has been provided in R-code in section 6.1 as comments.

2.2.2 T-tests

To see the difference between the regions in general and between each income level, Welch´s t test has been used.

H_0:$\mu_1 = \mu_2$ (two sided) where:

H_0 = the null hypothesis

μ_1 = the mean of sample 1, and μ_2 = the mean of sample 2.

<table>
<tr><td colspan="2">East and west in general

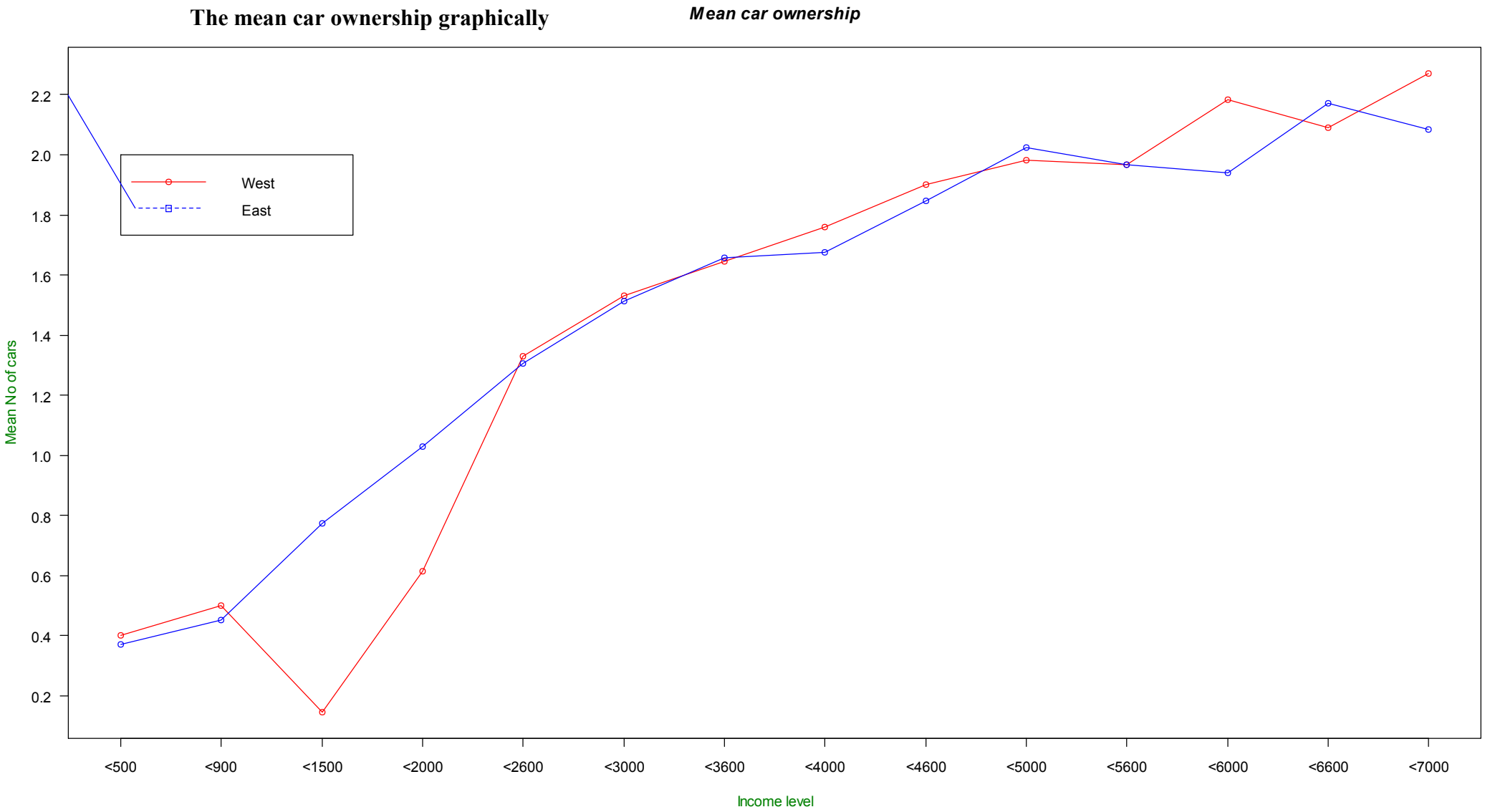
data: careast and carwest
t = -99.9418, df = 46975.99, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.7331039 -0.7049024
sample estimates:
mean of x mean of y
0.3173752 1.0363784</td></tr>
<tr><td>Income group <500
data: careast1 and carwest1
t = -0.4213, df = 194.73, p-value = 0.674
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.1659818 0.1075549
sample estimates:
mean of x mean of y
0.3707865 0.4000000</td><td>Income group <900
data: careast2 and carwest2
t = -1.2698, df = 871.545, p-value = 0.2045
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.11949509 0.02561604
sample estimates:
mean of x mean of y
0.4522145 0.4991540</td></tr>
<tr><td>Income group <1500
data: careast3 and carwest3
t = 0.1642, df = 2504.49, p-value = 0.8696
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.03932074 0.04650586
sample estimates:
mean of x mean of y
0.7757382 0.7721457</td><td>Income group <2000
data: careast4 and carwest4
t = -1.7093, df = 2485.809, p-value = 0.08752
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.07514629 0.00515153
sample estimates:
mean of x mean of y
1.029658 1.064655</td></tr>
<tr><td>Income group <2600</td><td>Income group <3000</td></tr>
</table>

data: careast5 and carwest5 t = -1.052, df = 2136.776, p-value = 0.2929 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.06889723 0.02078781 sample estimates: mean of x mean of y 1.306478 1.330533	data: careast6 and carwest6 t = -0.5665, df = 795.537, p-value = 0.5712 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.09197585 0.05077996 sample estimates: mean of x mean of y 1.511905 1.532503
Income group <3600 data: careast7 and carwest7 t = 0.2958, df = 875.596, p-value = 0.7674 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.05946528 0.08057074 sample estimates: mean of x mean of y 1.656028 1.645476	**Income group <4000** data: careast8 and carwest8 t = -1.3979, df = 287.115, p-value = 0.1632 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.20629176 0.03495089 sample estimates: mean of x mean of y 1.675127 1.760797
Income group <4600 data: careast9 and carwest9 t = -0.9629, df = 346.752, p-value = 0.3363 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.16353896 0.05603914 sample estimates: mean of x mean of y 1.846154 1.899904	**Income group <5000** data: careast10 and carwest10 t = 0.4063, df = 108.172, p-value = 0.6853 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.1664187 0.2522362 sample estimates: mean of x mean of y 2.024390 1.981481
Income group <5600 data: careast11 and carwest11 t = 0.008, df = 113.411, p-value = 0.9936 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.2203878 0.2221828 sample estimates: mean of x mean of y 1.967033 1.966135	**Income group <6000** data: careast12 and carwest12 t = -1.7264, df = 54.068, p-value = 0.08998 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.52394262 0.03909414 sample estimates: mean of x mean of y 1.939394 2.181818
Income group <6600 data: careast13 and carwest13 t = 0.4652, df = 51.829, p-value = 0.6437 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.2628736 0.4215413 sample estimates: mean of x mean of y 2.170732 2.091398	**Income group <7000** data: careast14 and carwest14 t = -1.0895, df = 19.923, p-value = 0.2889 alternative hypothesis: true difference in means is not equal to 0 95 percent confidence interval: -0.5495268 0.1725041 sample estimates: mean of x mean of y 2.083333 2.271845

Due to the unknown reason our T-test is not working properly, at least the given results significantly differ from the graphically plotted results.

2.2.3 Descriptive Analysis

The mean car ownership graphically



2.2.4 Conclusions

As seen from the graph, the mean number of cars per household is continuously growing index. It means that the car ownership depends on the income level.

The graph can be separated into three parts:

1. Income level till 900 euro per month;
2. Income level from 1000 till 3000 euro per month;
3. Income group for more than 3100 euro per month.

For the first group, the income level is too small to afford a car, that is why the mean car ownership is very low. However due to the low number of households in this group, the mean is very sensitive and can be increased by a single household which hystorically owns a car. Hystorically means that it can have purchased the car before the dicrease of the income level.

The second group is seperated, because their income is enough to afford a single till two private cars and the mean value line is very steap. It means that the mean number of cars is growing very considerably for each next higher income group.

In the third group the mean number of cars per household is stabilazing and the graph becomes less steep. It can e traslated into o urtr necesarity for a private car. And the decidion to increase the number of private cars in the household is becoming less dependent of income level.

To analyze if there is any significant difference of owning vehicles by households within a same range of income in east and West Germany, a T-test has been performed for constant 15 income ranges as shown in the graph. It has been observed that there is a significant change in the mean vehicle ownership per household for income ranges “€900-€1500”, “€1600-€2000”, “€5600-€6000”.

In the first two cases the number of cars per household in West Germany is well below that of East Germany. It is due to the fact that the public transport infrastructure is well developed in Western part of the country so that middle income earning people rely more on public transport instead of private transport.

But in third case, income level “€5600-€6000”, the mean number of car ownership per household is bigger for western Germany, what can be explained with historical importance of having a holiday or hobby car.

Nevertheless, there is no significant difference in other income ranges. The reasons may be discussed for two income levels - low and high. For low income groups, it is not possible to own and maintain a car. For high income groups, there is more than one car per household, what can be for grownups and holyday, hobby or children car.

2.3.1 Most frequent users of public transport

Macro picture illustrates that people always travel and are always mobile. But to understand travel behavior precisely, we are also interested in micro pictures of travelling. We want to investigate which particular groups of people are travelling most. Do young people travel more for education or recreational purposes? Or matured people for work and leisure or else senior citizens for passing their times? Does mobility is different for different sex groups or for people belonging different income groups etc. Above all, we are interested in how many people of above mentioned scenario consider public transport to fulfill their needs.

Municipalities often try to attract more people to the use of public transport. To do this, detailed surveys should be conducted in order to know the reasons why the people who avoid using public transport are doing so. Getting a random sample will result in a lot of wasted time since a certain percentage of the population is already using the public transport and there is no point in interviewing them for this particular study. On the other hand, a survey targeting the people who use the public transport might be needed to know if the consumers are satisfied with the service or if they require some improvements. Again, a random sample will result in wasted time due to people who do not use the service in it.

Having data about the age, employment, education level, employment and activity, we can figure out which groups are the most frequent users of public transport. There are many variables associated with our question but we took the most important, still to get the true big picture one need to include all those variables but due to time and resource limitation, we took only a few.

In our question, we took gender, age (10 year interval), Education level, Employment type and Activity type. Moreover, we took only those people in the groups which travel almost daily or at least 1 to 3 times a week. We neglect those traveler which travels one a month or twice a month.

2.3.2 Method

To performer this Procedure, we need to take the variables which we need for our calculation, we select Age, Gender, Education, Activity and Profession. After extracting the data which we need than we remove all the false data the missing values. Finally, filtered the data according to daily and 1 to 3 times a week public transport users after that we plot the data according to the groups.

No T-Test for performed for that question. Further explanation has been provided in R-code in section 6.1 as comments.

2.3.3 Descriptive Analysis:

We prepared some bar plots to show the travelling behavior of different groups mentioned above.

Gender:

Below Figure showed us that the females were more frequent public transport users than Male. There are no obvious reasons why females prefer more public transport than males. A deeper investigation is required to figure out the reason why females used more public transport than males.

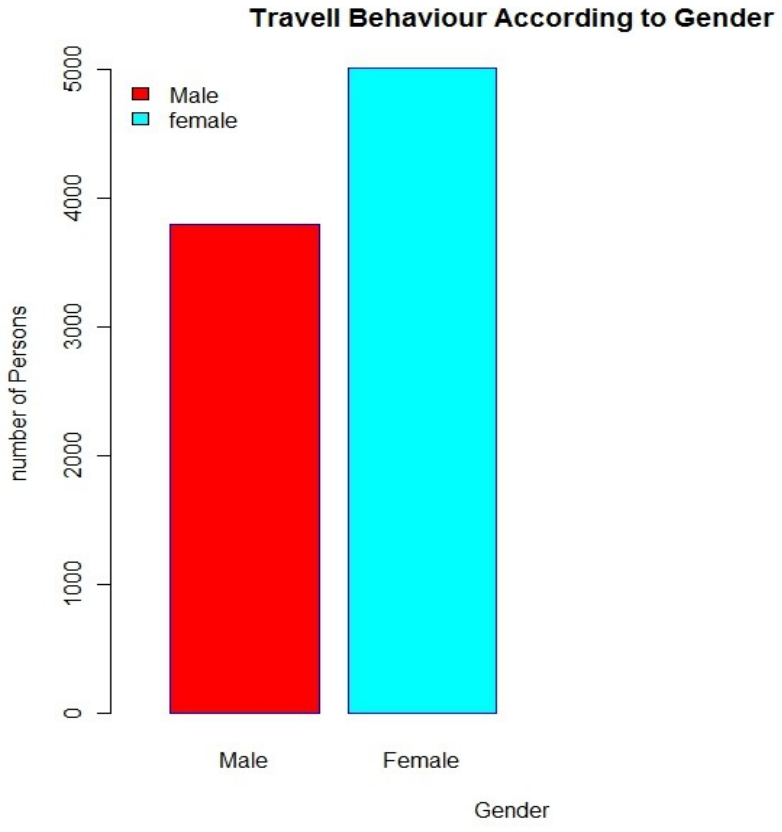


Fig 17:Gender of persons who uses public transport

Age:

From the figure below, it's quite clear that the most frequent users of public transport were the teenagers under 19.Probably at that time teenagers were studying and most of them don't have enough money, and they prefer going to their colleges through public transport. Then the most frequent group is from age 40 to 70 that travels a lot through public transport on a daily basis. The reasons may be they don't enjoy driving cars by themselves or may be some health issues.

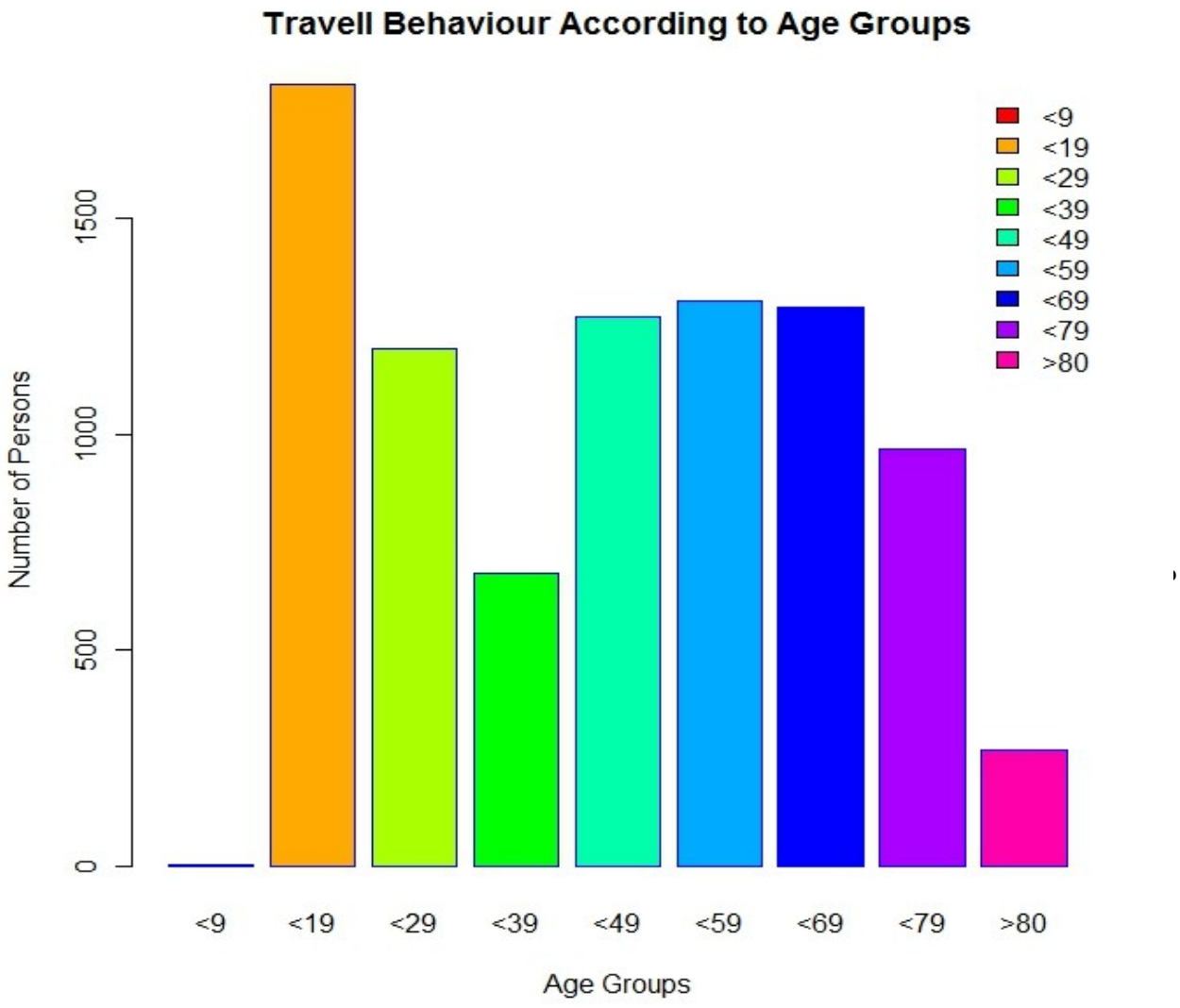


Fig 18: Age Groups who uses Public transport

Education level:

Most frequent users of transport are those who have university education or college education and then medium education group, but other groups or uneducated groups more likely to use private cars then public transport. So, it's really important to bring awareness to the society through schools ,kids should be taught about advantages of public transport, its economical and less pollutant .Moreover, awareness to those people who don't understand the importance of Global warming and jams and all the problems which are associated with private car use.

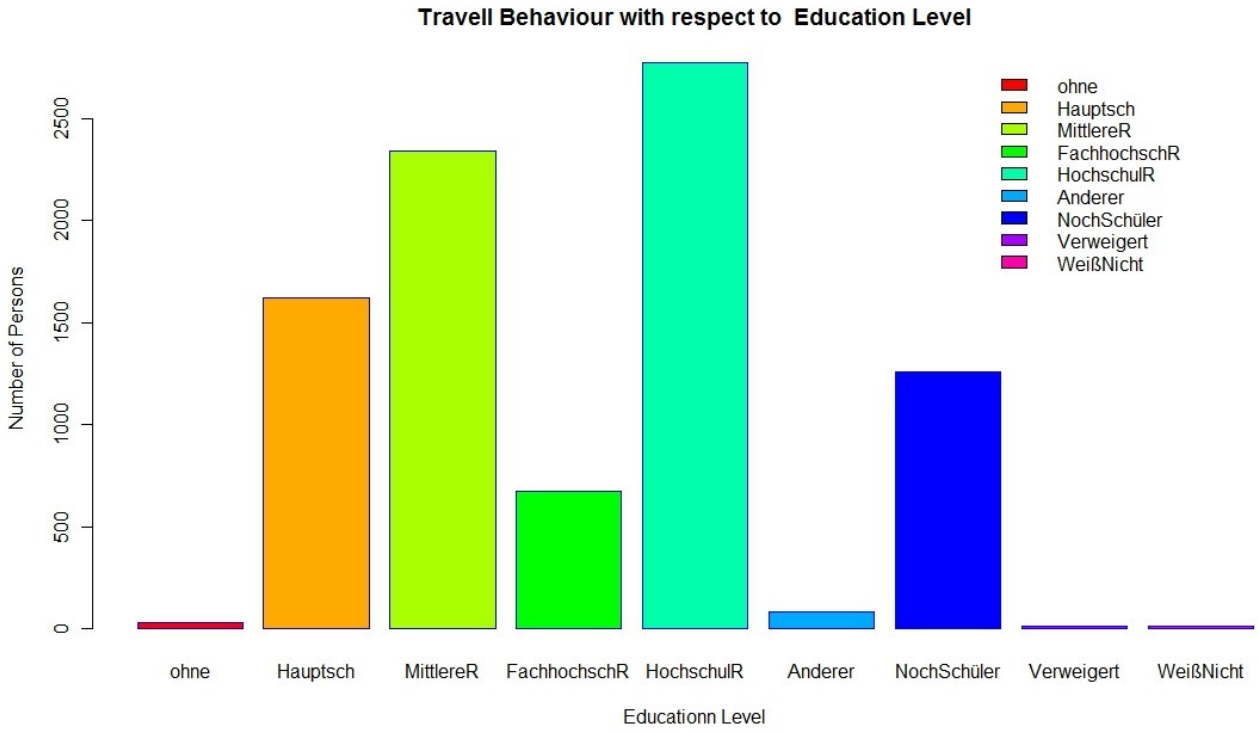


Fig 19: Education levels who uses Public transport

Employment level:

In the figure below, it's quite astonishing that, the most frequent users of public transport were not full time employee rather those who are not professional. There might be certain reasons behind may be they were not earning good enough to own a car or may be its easy to move through public transport rather using own vehicles. On the other side, full time employees should use public transport because of the nature of their job and public transport suits their daily. It's quite strange and further investigation is required to understand their travelling behavior.

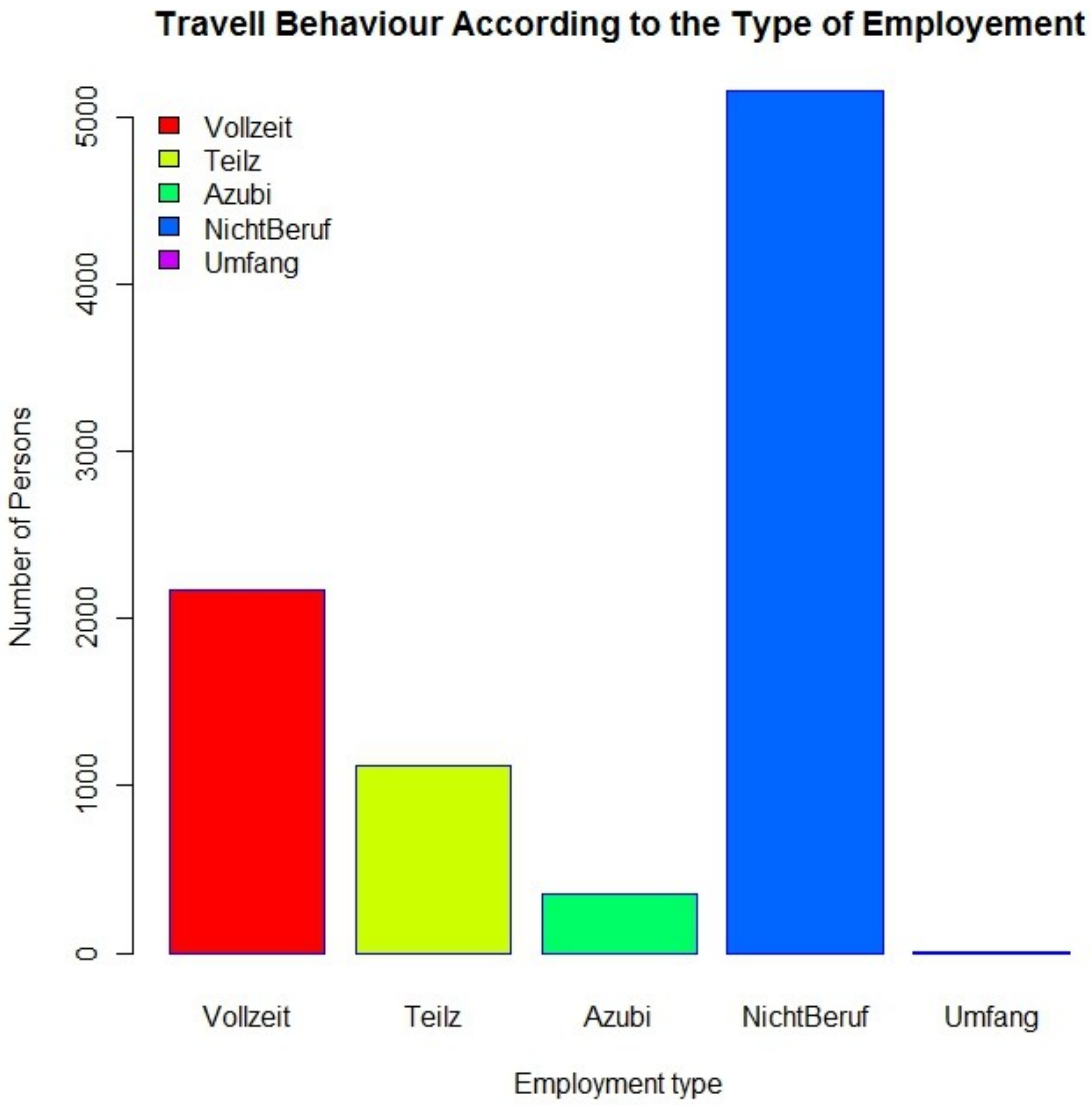


Fig 19:Employment Groups who uses Public transport

Activity type:

In this group of data, mostly public transport users were the professional people, and in second place there are retired people and then school going people. In this figure most probably the professional people may belong to not full time professional people as mentioned in the previous figure.

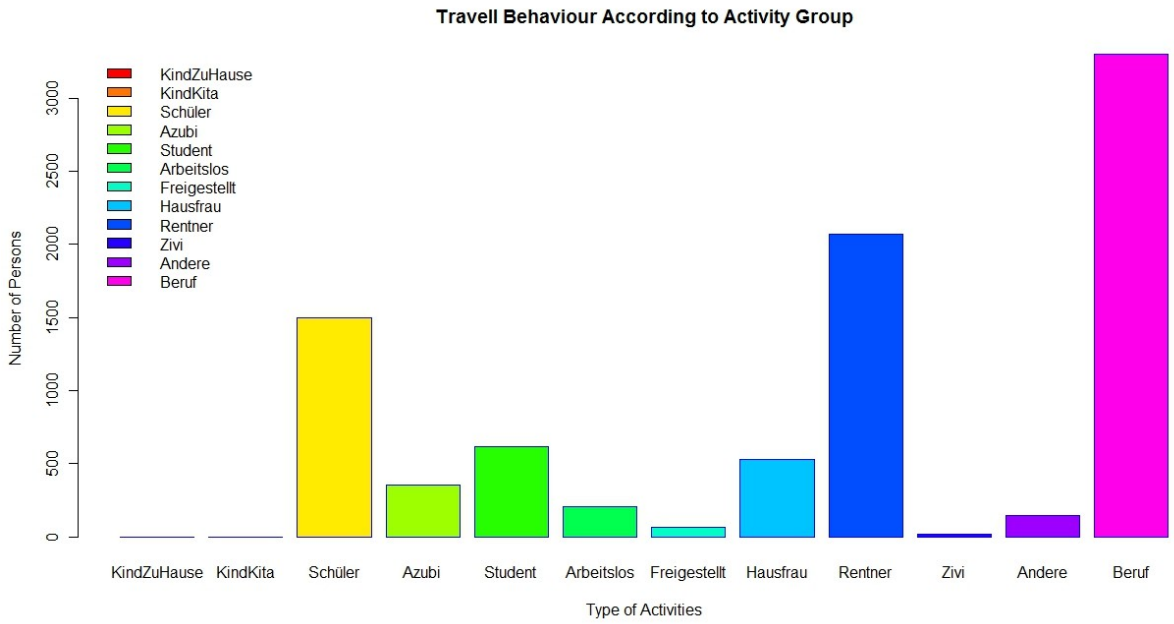


Fig 20:Activity Groups who uses Public transport

2.3.3 Conclusion:

- Female persons travel more than the male counterpart.
- Teenagers (age 13-19) and middle aged to senior citizens (age40-70) travel frequently.
- People with higher education are more frequent users of public transport.
- Freelance professionals use public transport more than full time employees.
- Among people with different activity, those with jobs are frequent users of public transport.

3 Main Conclusion

Our main aim for the research was based on the car ownership and the mean income level in all Germany and considering both parts of it. And after working with the data and analyzing it, our main conclusions are:

- Income has positive influence on car ownership.
- Though it is clear from the table/curve that the number of households with higher income (>2000) of west part is way higher than those of east, it has been observed that households belonging to west part are less reluctant to own or maintain higher number of cars or even a single one might seem enough. It can be caused by better public transport system, education, environment or view towards life is different there.
- Even though households are not interested to own more cars as income range crosses its peak, households in Germany realize it is better to have at least one car. It is beyond this project to check the difference between the average daily traffic in both parts that would certainly be helpful in observing the dependency on private cars
- Households having a higher income level have bigger mean car ownership
- The mean car ownership per household in whole Germany does not fluctuate that much.
- Use of public transport is different for different people. Sex, job type, education, age influence use of public transport.

4 Problems during the work

During the work the biggest problems for us was to deal with the “NA” and “K.A.” values in the data. Because there was some cases when “NA” values where necessary, but in other cases they were considered as mistakes.

For the research of the most frequent public transport user in the Germany, there was a problem with deleting “K.A.” value from the given data.

5 References

1) A Beginner's Guide to R, Alain F. Zuur, Elena N. Ieno, Erik H. W. G. Meesters, 2009

6 Appendices

6.1 R-code

```
Quetion1:
#Setting the working directorie
setwd("C:\\Users\\stat project\\Data")
#reading household data
load("H2008.Rdata")
attach(h2008)
#Creating variables for further use
income1=0; income2=0; income3=0; income4=0; income5=0; income6=0; income7=0;
income8=0; income9=0; income10=0; income11=0; income12=0; income13=0;
income14=0; income15=0

#Dividing data of car ownage by each income level
r<-dim(h2008)[1]
i=0
for (i in 1:r) { if (hheink[i]=="<500") {income1[i]=h04_3[i]}
                 if (hheink[i]=="<900") {income2[i]=h04_3[i]}
                 if (hheink[i]=="<1500") {income3[i]=h04_3[i]}
                 if (hheink[i]=="<2000") {income4[i]=h04_3[i]}
                 if (hheink[i]=="<2600") {income5[i]=h04_3[i]}
                 if (hheink[i]=="<3000") {income6[i]=h04_3[i]}
                 if (hheink[i]=="<3600") {income7[i]=h04_3[i]}
                 if (hheink[i]=="<4000") {income8[i]=h04_3[i]}
                 if (hheink[i]=="<4600") {income9[i]=h04_3[i]}
                 if (hheink[i]=="<5000") {income10[i]=h04_3[i]}
                 if (hheink[i]=="<5600") {income11[i]=h04_3[i]}
                 if (hheink[i]=="<6000") {income12[i]=h04_3[i]}
                 if (hheink[i]=="<6600") {income13[i]=h04_3[i]}
                 if (hheink[i]=="<7000") {income14[i]=h04_3[i]}
                 if (hheink[i]==">7000") {income15[i]=h04_3[i]}

               }
#Replacing data which we do not need with the NA
income1<-replace(income1,income1=="98"|income1=="99"|income1=="97",NA)
income2<-replace(income2,income2=="98"|income2=="99"|income2=="97",NA)
income3<-replace(income3,income3=="98"|income3=="99"|income3=="97",NA)
income4<-replace(income4,income4=="98"|income4=="99"|income4=="97",NA)
income5<-replace(income5,income5=="98"|income5=="99"|income5=="97",NA)
income6<-replace(income6,income6=="98"|income6=="99"|income6=="97",NA)
income7<-replace(income7,income7=="98"|income7=="99"|income7=="97",NA)
income8<-replace(income8,income8=="98"|income8=="99"|income8=="97",NA)
income9<-replace(income9,income9=="98"|income9=="99"|income9=="97",NA)
income10<-replace(income10,income10=="98"|income10=="99"|income10=="97",NA)
income11<-replace(income11,income11=="98"|income11=="99"|income11=="97",NA)
income12<-replace(income12,income12=="98"|income12=="99"|income12=="97",NA)
income13<-replace(income13,income13=="98"|income13=="99"|income13=="97",NA)
income14<-replace(income14,income14=="98"|income14=="99"|income14=="97",NA)
income15<-replace(income15,income15=="98"|income15=="99"|income15=="97",NA)

#Ploting the graphs for the results
```

```
barplot(table(income1), main="No of cars with income group <500",
xlab="Number of Cars", ylab="Number of Households", border="blue",
density=c(10,20,30))
barplot(table(income2), main="No of cars with income group <900",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three"), border="blue",
density=c(10,20,30,40))
barplot(table(income3), main="No of cars with income group <1500",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","seven"),border="blue",
density=c(10,20,30,40,45,50,55))
barplot(table(income4), main="No of cars with income group <2000",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five"), border="blue",
density=c(10,20,30,40,45,50))
barplot(table(income5), main="No of cars with income group <2600",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six"), border="blue",
density=c(10,20,30,40,45,50,55))
barplot(table(income6),main="No of cars with income group <3000",
xlab="Number of Cars",ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six","seven","eight"),b
order="blue", density=c(10,20,30,40,45,50,55,60,70))
barplot(table(income7), main="No of cars with income group <3600",
xlab="Number of Cars",ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","seven","eight"),
border="blue", density=c(10,20,30,40,45,50,55,60))
barplot(table(income8), main="No of cars with income group <4000",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six"),
border="blue", density=c(10,20,30,40,45,50,55))
barplot(table(income9), main="No of cars with income group <4600",
xlab="Number of Cars",ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six"),
border="blue", density=c(10,20,30,40,45,50,55))
barplot(table(income10), main="No of cars with income group <15000",
xlab="Number of Cars",ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six"),
border="blue", density=c(10,20,30,40,45,50,55))
barplot(table(income11), main="No of cars with income group <5600",
xlab="Number of Cars",ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six","seven"),
border="blue", density=c(10,20,30,40,45,50,55,60))
barplot(table(income12), main="No of cars with income group <6000",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six"),
border="blue", density=c(10,20,30,40,45,50,55))
barplot(table(income13), main="No of cars with income group <6600",
xlab="Number of Cars",ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six"),
border="blue", density=c(10,20,30,40,45,50,55))
barplot(table(income14), main="No of cars with income group <7000",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six"),
border="blue", density=c(10,20,30,40,45,50,55))
```

```
barplot(table(income15), main="No of cars with income group <7500",
xlab="Number of Cars", ylab="Number of Households",
names.arg=c("ZERO","ONE","TWO","three","four","five","six","eight"),
border="blue", density=c(10,20,30,40,45,50,55,60))



#The levels "Nein", "Weiß nicht", "k.A." replaced by missing values for No of
cars
h2008$h04_3<-replace(h2008$h04_3,h2008$h04_3=="98"|h2008$h04_3=="99"
|h2008$h04_3== "97",NA)
#The levels "Nein", "Weiß nicht", "k.A." replaced by missing values  for
household income incomelevel
h2008$hheink<-replace(h2008$hheink,h2008$hheink=="WeißNicht"|
h2008$hheink=="k.A."|h2008$hheink=="Verweigert",NA)
#Conversion of the income level in to numeric data to plot
mainplot = c(h2008$hheink)
fdata = factor(mainplot)
rdata = factor(data,labels=c('500','900','1500','2000','2600','3000',
'3600','4000','4600','5000','5600','6000','6600','7000','7500'))
#Plotting of the main graph "Car ownership according to different income
levels of household income"
mainplot2 <- table(h2008$h04_3,rdata )
barplot(mainplot2, main="Car ownership according to different income levels
of household income",xlab="Household Income level", ylab="Number of
Households", col=rainbow(9))
legend("topright", c("0 car","1 car","2 cars","3 cars","4 cars","5 cars","6
cars","7 cars","8 cars"), cex= 1.0, bty="n", fill=rainbow(9));
#T-test from various Income Levels compared to the whole Sample
t.test(income10,income15,"two.sided")
t.test(income2,income8,"two.sided")
t.test(income5,income7,"two.sided")
t.test(income12,income13,"two.sided")
```

```
Question2:

#loading data
load("H2008.Rdata")
attach(h2008)
r<-dim(h2008)[1]
#Replacing NA with unused data groups
a=0
for (a in 1:r){if (is.na(hheink[a])){hheink[a]=">7000"}}
a=0;
for (a in 1:r){if (is.na(h04_3[a])){h04_3[a]="h04_3[a-1]"}}
#Creating variables for further use
carwest1<-c(); carwest2<-c(); carwest3<-c(); carwest4<-c(); carwest5<-c();
carwest6<-c(); carwest7<-c(); carwest8<-c(); carwest9<-c(); carwest10<-c();
carwest11<-c(); carwest12<-c(); carwest13<-c(); carwest14<-c()

# nrhw - how many households are in this income group, west
nrhw1=0; nrhw2=0; nrhw3=0; nrhw4=0; nrhw5=0; nrhw6=0; nrhw7=0; nrhw8=0;
nrhw9=0; nrhw10=0; nrhw11=0; nrhw12=0; nrhw13=0; nrhw14=0;
# nrcw - how many cars are in this income group, west
nrcw1=0; nrcw2=0; nrcw3=0; nrcw4=0; nrcw5=0; nrcw6=0; nrcw7=0; nrcw8=0;
nrcw9=0; nrcw10=0; nrcw11=0; nrcw12=0; nrcw13=0; nrcw14=0
# Counting how much cars are there in each income group and how much
households are there in each income group
#Seperating the data of west part of Germany
 i=0
 j=0
 for( i in 1:r){
      { if (westost[i]=="West"& hheink[i]=="<500")
{carwest1[i]=h2008[i,"h04_3"];nrhw1=nrhw1+1;nrcw1=nrcw1+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<900")
{carwest2[i]=h2008[i,"h04_3"];nrhw2=nrhw2+1;nrcw2=nrcw2+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<1500")
{carwest3[i]=h2008[i,"h04_3"];nrhw3=nrhw3+1;nrcw3=nrcw3+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<2000")
{carwest4[i]=h2008[i,"h04_3"];nrhw4=nrhw4+1;nrcw4=nrcw4+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<2600")
{carwest5[i]=h2008[i,"h04_3"];nrhw5=nrhw5+1;nrcw5=nrcw5+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<3000")
{carwest6[i]=h2008[i,"h04_3"];nrhw6=nrhw6+1;nrcw6=nrcw6+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<3600")
{carwest7[i]=h2008[i,"h04_3"];nrhw7=nrhw7+1;nrcw7=nrcw7+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<4000")
{carwest8[i]=h2008[i,"h04_3"];nrhw8=nrhw8+1;nrcw8=nrcw8+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<4600")
{carwest9[i]=h2008[i,"h04_3"];nrhw9=nrhw9+1;nrcw9=nrcw9+h2008[i,"h04_3"]}}
      { if (westost[i]=="West"& hheink[i]=="<5000")
{carwest10[i]=h2008[i,"h04_3"];nrhw10=nrhw10+1;nrcw10=nrcw10+h2008[i,"h04_3"]
}}
      { if (westost[i]=="West"& hheink[i]=="<5600")
{carwest11[i]=h2008[i,"h04_3"];nrhw11=nrhw11+1;nrcw11=nrcw11+h2008[i,"h04_3"]
}}
      { if (westost[i]=="West"& hheink[i]=="<6000")
{carwest12[i]=h2008[i,"h04_3"];nrhw12=nrhw12+1;nrcw12=nrcw12+h2008[i,"h04_3"]
}}
```

```
      { if (westost[i]=="West"& hheink[i]=="<6600")
{carwest13[i]=h2008[i,"h04_3"];nrhw13=nrhw13+1;nrcw13=nrcw13+h2008[i,"h04_3"]
}}
      { if (westost[i]=="West"& hheink[i]=="<7000")
{carwest14[i]=h2008[i,"h04_3"];nrhw14=nrhw14+1;nrcw14=nrcw14+h2008[i,"h04_3"]
}}
      }

 nrcw3=295      #we use this due to a mistake in data file
 nrcw4=1569     #we use this due to a mistake in data file
#Meancw - will store the mean No of cars in the househols in every income
group in west part of germnany
 meancw=0

meancw[1]=nrcw1/nrhw1; meancw[2]=nrcw2/nrhw2; meancw[3]=nrcw3/nrhw3;
meancw[4]=nrcw4/nrhw4; meancw[5]=nrcw5/nrhw5; meancw[6]=nrcw6/nrhw6;
meancw[7]=nrcw7/nrhw7; meancw[8]=nrcw8/nrhw8; meancw[9]=nrcw9/nrhw9;
meancw[10]=nrcw10/nrhw10; meancw[11]=nrcw11/nrhw11; meancw[12]=nrcw12/nrhw12;
meancw[13]=nrcw13/nrhw13; meancw[14]=nrcw14/nrhw14
#Creating variables for further use
careast1<-c(); careast2<-c(); careast3<-c(); careast4<-c(); careast5<-c();
careast6<-c(); careast7<-c(); careast8<-c(); careast9<-c(); careast10<-c();
careast11<-c(); careast12<-c(); careast13<-c(); careast14<-c()
# nrhw - how many households are in this income group, west
nrhe1=0; nrhe2=0; nrhe3=0; nrhe4=0; nrhe5=0; nrhe6=0; nrhe7=0; nrhe8=0;
nrhe9=0; nrhe10=0; nrhe11=0; nrhe12=0; nrhe13=0; nrhe14=0
# nrcw - how many cars are in this income group, west
nrce1=0; nrce2=0; nrce3=0; nrce4=0; nrce5=0; nrce6=0; nrce7=0; nrce8=0;
nrce9=0; nrce10=0; nrce11=0; nrce12=0; nrce13=0; nrce14=0
# Counting how much cars are there in each income group and how much
households are there in each income group
#Seperating the data of east part of Germany
 k=0
 l=0
 for( k in 1:r){
      { if (westost[k]=="Ost"& hheink[k]=="<500")
{careast1[k]=h2008[k,"h04_3"];nrhe1=nrhe1+1;nrce1=nrce1+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<900")
{careast2[k]=h2008[k,"h04_3"];nrhe2=nrhe2+1;nrce2=nrce2+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<1500")
{careast3[k]=h2008[k,"h04_3"];nrhe3=nrhe3+1;nrce3=nrce3+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<2000")
{careast4[k]=h2008[k,"h04_3"];nrhe4=nrhe4+1;nrce4=nrce4+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<2600")
{careast5[k]=h2008[k,"h04_3"];nrhe5=nrhe5+1;nrce5=nrce5+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<3000")
{careast6[k]=h2008[k,"h04_3"];nrhe6=nrhe6+1;nrce6=nrce6+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<3600")
{careast7[k]=h2008[k,"h04_3"];nrhe7=nrhe7+1;nrce7=nrce7+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<4000")
{careast8[k]=h2008[k,"h04_3"];nrhe8=nrhe8+1;nrce8=nrce8+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<4600")
{careast9[k]=h2008[k,"h04_3"];nrhe9=nrhe9+1;nrce9=nrce9+h2008[k,"h04_3"]}}
      { if (westost[k]=="Ost"& hheink[k]=="<5000")
{careast10[k]=h2008[k,"h04_3"];nrhe10=nrhe10+1;nrce10=nrce10+h2008[k,"h04_3"]
}}
```

```
      { if (westost[k]=="Ost"& hheink[k]=="<5600")
{careast11[k]=h2008[k,"h04_3"];nrhe11=nrhe11+1;nrce11=nrce11+h2008[k,"h04_3"]
}}
      { if (westost[k]=="Ost"& hheink[k]=="<6000")
{careast12[k]=h2008[k,"h04_3"];nrhe12=nrhe12+1;nrce12=nrce12+h2008[k,"h04_3"]
}}
      { if (westost[k]=="Ost"& hheink[k]=="<6600")
{careast13[k]=h2008[k,"h04_3"];nrhe13=nrhe13+1;nrce13=nrce13+h2008[k,"h04_3"]
}}
      { if (westost[k]=="Ost"& hheink[k]=="<7000")
{careast14[k]=h2008[k,"h04_3"];nrhe14=nrhe14+1;nrce14=nrce14+h2008[k,"h04_3"]
}}
      }
```

```
careast=0
l=0
k=0
for( k in 1:r){if (westost[k]=="Ost"&
hheink[i]=="<500"|westost[k]=="Ost"&hheink[i]=="<900"|westost[k]=="Ost"&hhein
k[i]=="<1500"|westost[k]=="Ost"&hheink[i]=="<2000"|

westost[k]=="Ost"&hheink[i]=="<2600"|westost[k]=="Ost"&hheink[i]=="<3000"|wes
tost[k]=="Ost"& hheink[i]=="<3600"|westost[k]=="Ost"& hheink[i]=="<4000"|
                westost[k]=="Ost"& hheink[i]=="<4600"|westost[k]=="Ost"&
hheink[i]=="<5000"| westost[k]=="Ost"& hheink[i]=="<5600"|westost[k]=="Ost"&
hheink[i]=="<6000"|
                westost[k]=="Ost"& hheink[i]=="<6600"|westost[k]=="Ost"&
hheink[i]=="<7000"){careast[k]=h2008[k,"h04_3"];l=l+1}}
carwest=0
l=0
k=0
for( k in 1:r){if (westost[k]=="West"& hheink[i]=="<500"|westost[k]=="West"&
hheink[i]=="<900"|westost[k]=="West"& hheink[i]=="<1500"|westost[k]=="West"&
hheink[i]=="<2000"|
                  westost[k]=="West"& hheink[i]=="<2600"|westost[k]=="West"&
hheink[i]=="<3000"|westost[k]=="West"& hheink[i]=="<3600"|westost[k]=="West"&
hheink[i]=="<4000"|
                  westost[k]=="West"& hheink[i]=="<4600"|westost[k]=="West"&
hheink[i]=="<5000"|westost[k]=="West"& hheink[i]=="<5600"|westost[k]=="West"&
hheink[i]=="<6000"|
                  westost[k]=="West"& hheink[i]=="<6600"|westost[k]=="West"&
hheink[i]=="<7000"){carwest[k]=h2008[k,"h04_3"];l=l+1}}
#Replacing NA's with a 0
 a=0;
 for (a in 1:r){if (is.na(careast[a])){careast[a]=0}}
 a=0;
 for (a in 1:r){if (is.na(carwest[a])){carwest[a]=0}}

 nrce2=194  #we use this due to a mistake in data file
 nrce3=972  #we use this due to a mistake in data file
 nrce5=1573 #we use this due to a mistake in data file

#Calculating the mean No of cars in each income group in east side of Germany
```

```
 meance=0
meance[1]=nrce1/nrhe1; meance[2]=nrce2/nrhe2; meance[3]=nrce3/nrhe3;
meance[4]=nrce4/nrhe4; meance[5]=nrce5/nrhe5; meance[6]=nrce6/nrhe6;
meance[7]=nrce7/nrhe7; meance[8]=nrce8/nrhe8; meance[9]=nrce9/nrhe9;
meance[10]=nrce10/nrhe10; meance[11]=nrce11/nrhe11; meance[12]=nrce12/nrhe12;
meance[13]=nrce13/nrhe13; meance[14]=nrce14/nrhe14
#Variable for y axis
income<-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14)
#Ploting the results
plot(income,meancw,type="o",col="red",axes=FALSE, ann=FALSE)
axis(1, at=1:14,
lab=c("<500","<900","<1500","<2000","<2600","<3000","<3600","<4000","<4600","
<5000","<5600","<6000","<6600","<7000"))
axis(2, las=1, at=0.2*1:14)
lines(income,meance,type="o",col="blue"); box()
title(main="Mean car ownership", col.main="black", font.main=4)
title(xlab="Income level", col.lab=rgb(0,0.5,0))
title(ylab="Mean No of cars", col.lab=rgb(0,0.5,0))
legend(1, 2, c("West","East"), cex=1,col=c("red","blue"), pch=21:22,
lty=1:2);
#Calculating t.test for east and west part of Germany
t.test(careast,carwest)
#Calculating t.test for each income group bitwean east and west part of
Germany
 t.test(careast1,carwest1,"two.sided")
 t.test(careast2,carwest2,"two.sided")
 t.test(careast3,carwest3,"two.sided")
 t.test(careast4,carwest4,"two.sided")
 t.test(careast5,carwest5,"two.sided")
 t.test(careast6,carwest6,"two.sided")
 t.test(careast7,carwest7,"two.sided")
 t.test(careast8,carwest8,"two.sided")
 t.test(careast9,carwest9,"two.sided")
 t.test(careast10,carwest10,"two.sided")
 t.test(careast11,carwest11,"two.sided")
 t.test(careast12,carwest12,"two.sided")
 t.test(careast13,carwest13,"two.sided")
 t.test(careast14,carwest14,"two.sided")
```

Qusetion3:

```
#Loading data
load("P2008.rdata")
attach(p2008)
r<-dim(p2008)[1]
#taking only data of people which use public transport every day or 1 till 3
times a week
i=0
a=0
b=0
for (i in 1:r) { if (p032[i]=="Tägl"|p032[i]=="1bis3Woche") {a=a+1}
                 else {p032[i]<-0;hp_sex[i]<-0;hp_altg2[i]<-0;p14a[i]<-
                       0;hp_bkat[i]<-0;hp_taet[i]<-0}
                 if (is.na(hp_sex[i])|is.na(hp_altg2[i])|is.na(p14a[i])|
                     is.na(hp_bkat[i])|is.na(hp_taet[i]))
                    {b=b+1;p032[i]<-0;hp_sex[i]<-0;hp_altg2[i]<-0;
                     p14a[i]<-0; hp_bkat[i]<-0; hp_taet[i]<-0}
               }

#Creating variables for further use
#gender
sex=c(0,0)
#age group
age=c(0,0,0,0,0,0,0,0,0)
#highest level of education
education=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)
#type of employement
employement=c(0,0,0,0,0)
#type of activity
activity=c(0,0,0,0,0,0,0,0,0,0,0,0,0)
i=0
#Writing tables of data content into the variables
sex1=c(table(hp_sex))
age1=c(table(hp_altg2))
education1=c(table(p14a))
employement1=c(table(hp_bkat))
activity1=c(table(hp_taet))
#Creating variables for further use
sex2=0; age2=0; education2=0; employement2=0; activity2=0;
#Writing new variables from previous ones without the last value, because it
is NA
a=length(sex1)-1
i=0
for (i in 1:a) { sex2[i]=c(sex1[i]) }
a=length(age1)-1
i=0
for (i in 1:a) { age2[i]=c(age1[i]) }
a=length(education1)-1
i=0
for (i in 1:a) { education2[i]=c(education1[i]) }
a=length(employement1)-1
i=0
for (i in 1:a) { employement2[i]=c(employement1[i]) }
```

```
a=length(activity1)-1
i=0
for (i in 1:a) { activity2[i]=c(activity1[i]) }

#Plot gender graph
barplot(sex2, main="Travell Behaviour According to Gender", xlab= "Gender",
ylab="number of Persons", names.arg=c("Male","Female"),
border="blue", col=rainbow(2),xlim=c(0,5))
legend("topleft", c("Male","female"), cex= 1.0,
bty="n", fill=rainbow(2));
#Plot age graph
barplot(age2, main="Travell Behaviour According to Age Groups", xlab= "Age
Groups",ylab="Number of Persons",
names.arg=c("<9","<19","<29","<39","<49","<59","<69","<79",">80"),
border="blue", col=rainbow(9))
legend("topright", c("<9","<19","<29","<39","<49","<59","<69","<79",">80"),
cex= 1.0,bty="n", fill=rainbow(9));
#Plot highest level of education graph
barplot(education, main="Travell Behaviour with respect to  Education Level",
xlab= "Educationn Level",ylab="Number of Persons",
names.arg=c("ohne","Hauptsch","MittlereR","FachhochschR","HochschulR","Andere
r","NochSchüler","Verweigert","WeißNicht"),
border="blue", col=rainbow(9))
legend("topright",c("ohne","Hauptsch","MittlereR","FachhochschR","HochschulR"
,"Anderer","NochSchüler","Verweigert","WeißNicht"), cex= 1.0, bty="n",
fill=rainbow(9));
#Plot type of employement graph
barplot(employement2, main="Travell Behaviour According to the Type of
Employement", xlab= "Employment type", ylab="Number of Persons",
names.arg=c("Vollzeit","Teilz","Azubi","NichtBeruf","Umfang"),
border="blue", col=rainbow(5))
legend("topleft", c("Vollzeit","Teilz","Azubi","NichtBeruf","Umfang"),
cex=1.0,bty="n", fill=rainbow(5));
#Plot type of activity graph
barplot(activity, main="Travell Behaviour According to Activity Group", xlab=
"Type of Activities",ylab="Number of Persons",
names.arg=c("KindZuHause","KindKita","Schüler","Azubi","Student","Arbeitslos"
,"Freigestellt","Hausfrau","Rentner","Zivi","Andere","Beruf"),
border="blue", col=rainbow(13))
legend("topleft",c("KindZuHause","KindKita","Schüler","Azubi","Student","Arbe
itslos","Freigestellt","Hausfrau","Rentner","Zivi","Andere","Beruf"), cex=
1.0,bty="n", fill=rainbow(13));
```