Question1:

library(XML)

#Find total unique new items in your data-set of new mobiles and tablets?

chundi <- heisenberg <- read.csv(file="olxnewfirstpage.csv", header = T,sep=",")

uni <- length( unique(chundi$category.............))

uni

# Output:

686

Question2:

olx <- heisenberg <- read.csv(file="olx22.csv", header = T,sep=",")

daraz <- heisenberg <- read.csv(file="assgnment2\_data\_2.csv", header = T,sep=",")

nomatch <- grep(daraz,olx$title, value = FALSE)

nomatch

# Output:

c(12, 23, 34, 36, 37, 38, 39, 40, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 1, 12, 23, 34, 36, 37, 38, 39, 40, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 1, 12, 23, 34, 36, 37, 38, 39, 40, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 1, 12, 23, 34, 36, 37, 38, 39, 40, 2, 3, 4, 5, 6,

7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 1, 12, 23, 34, 36, 37, 38, 39, 40, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 1, 12, 23, 34, 36, 37, 38, 39, 40, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 1, 12, 23, 34, 36, 37, 38, 39, 40, 2, 3, 4, 5, 6, 7, 8,

Question3:

library(sqldf)

data1 <- read.csv(file="detail.csv",

header = TRUE, sep = ","

)

data1$

minviews <- 422

maxviews <- 300

mer <- sqldf(paste("select detaildesc........ from data1 where views 100 ", sep = ""))

print(mer)

Question4:

library(sqldf)

#Which ad was viewed the most? list top 10 ads

data1 <- read.csv(file="detail.csv",

header = TRUE, sep = ","

)

tail(sort(data1$views),10)

# Output

[1] 996 996 997

[4] 999 999 999

[7] A5 chip Depth 9mm EDGE

[10] with out Sim.\t\t\t\t\t\t\t\t\t"

1871 Levels: brand New box pack dual core Made in vietnam ... with out Sim.\t\t\t\t\t\t\t\t\t"

Question5:

library(sqldf)

#What mobile was available for sale most ? list top 10 mobiles and top 10 tablets

data1 <- read.csv(file="olxnewfirstpage.csv",

header = TRUE, sep = ","

)

upper <- toupper(data1$title)

nn <- factor(upper)

nn

Question6:

olx <- heisenberg <- read.csv(file="olxnewfirstpage.csv", header = T,sep=",")

duplicate <- unique(olx$city)

duplicate

Question7:

#Which city included most ads? rank cities according to the number of ads posted

library(sqldf)

#What mobile was available for sale most ? list top 10 mobiles and top 10 tablets

data1 <- read.csv(file="olxnewfirstpage.csv",

header = TRUE, sep = ","

)

occur <- data.frame(table(data1$city))

occur[occur$Freq > 1,]

# Output:

2 \n\t\t\t\t\t\t\t\t\t\t\t\t\t\t\tTablets » Other Tablets\nKarachi\t\t\t\t\t\t\t\t\t\t\t\t\t\t 2

7 Abbottabad 122

8 Ahmadpur East 11

9 Arifwala 15

10 Attock 70

11 Badin 2

12 Bagh 11

13 Bahawalnagar 23

14 Bahawalpur 133

15 Bannu 6

16 Bhakkar 12

17 Bhimber 5

19 Burewala 36

20 Charsadda 7

21 Chiniot 15

22 Chishtian Mandi 5

24 Daska 22

25 Dera Ghazi Khan 35

26 Dera Ismail Khan 46

27 Faisalabad 698

28 Gilgit 3

29 Gojra 18

30 Gujranwala 366

31 Gujrat 189

33 Hafizabad 19

34 Hala 2

36 Haripur 34

37 Hasilpur 3

38 Haveli lakha 3

40 Hyderabad 267

41 Islamabad 1872

42 Jacobabad 11

43 Jamshoro 5

44 Jaranwala 20

45 Jhang Sadar 19

46 Jhelum 80

47 Kamoke 8

48 Karachi 3936

49 Karak 2

50 Kasur 25

51 Khairpur 14

52 Khanewal 42

53 Khanpur 4

54 Khushab 19

55 Kohat 13

56 Lahore 5785

57 Larkana 20

58 Lower Dir 2

59 Mansehra 25

60 Mardan 36

61 Mianwali 26

62 Mingaora 7

63 Mirpur 45

64 Mirpur Khas 17

65 Mithi 2

66 Multan 731

67 Muridike 14

68 Muzaffarabad 30

69 Muzaffargarh 20

70 Nawabshah 2

71 Nowshera 24

72 Okara 83

73 Pakpattan 10

74 Peshawar 476

75 Pirmahal 13

76 Quetta 106

77 Rahimyar Khan 50

79 Rawalpindi 2152

80 Sadiqabad 30

82 Sahiwal 151

83 Sargodha 192

84 Sheikhüpura 62

85 Shikarpur 4

86 Sialkot 290

87 Sukkar 2

88 Sukkur 43

89 Swabi 11

90 Swat 2

92 Thatta 3

93 Toba Tek singh 31

95 Wah 153

Question8:

olx <- heisenberg <- read.csv(file="olxnewfirstpage.csv", header = T,sep=",")

duplicate <- unique(olx$city)

duplicate

Question16:

If you want to buy a mobile/tablet today and possibly will sell it exactly 1 year later. Which

Mobile/tablet will you buy keeping in mind that you want to minimize the risk.

Answer:

I will buy IPhone because as I have seen that all other mobile brand except iPhone their price fluctuate and flow gradually on the other hand iPhone retain its price in long term so there would be a better chance if I want to sale my phone I would definitely get good money back from it.