

Marketing Analytics - Exercise Sheet 1

Read the dataset mvtWeek1.csv and answer the following questions using python/R

1. How many rows of data are there in dataset?
2. How many variables are there in the dataset?
3. What is the minimum value of the variable "Beat"?
4. How many observations have value TRUE in the Arrest variable?
5. How many observations have a LocationDescription value of ALLEY
6. In what format are the entries in the variable Date?
7. Create an histogram of variable Date. Crime increases are decreases
in 2002-2012, 2005-2008,2009-2011
8. Create a boxplot of the variable "Date", sorted by the variable "Arrest" . Does it look like there were more crimes for which arrests were made in the first half of the time period or the second half of the time period?
9. For what proportion of motor vehicle thefts in 2001 was an arrest made? Similarly for the year 2007, 2012
10. Which locations are the top five locations for motor vehicle thefts, excluding the "Other" category