

14.1

Problem- prediction of the number of comments in the upcoming 24 hours on those blogs, The train data was generated from different base times that may temporally overlap. Therefore, if you simply split the train into disjoint partitions, the underlying time intervals may overlap. Therefore, the you should use the provided, temporally disjoint train and test splits to ensure that the evaluation is fair.

- Read the dataset and identify the right features

Ans: topic not covered in class

- Clean dataset, impute missing values and perform exploratory data analysis.

No missing values.

```
> summary(blogData_train)
```

A		B		C		D	
Min. :	0.000	Min. :	0.000	Min. :	0.0000	Min. :	0.0
1st Qu.:	2.286	1st Qu.:	5.214	1st Qu.:	0.0000	1st Qu.:	29.0
Median :	10.631	Median :	19.353	Median :	0.0000	Median :	162.0
Max. :	4.00	Max. :	4.1507	Max. :	11.051	Max. :	0.000000

E		F		G		H	
Min. :	0.00	Min. :	0.0000	Min. :	0.000	Min. :	0.000000
1st Qu.:	0.00	1st Qu.:	0.8916	1st Qu.:	3.075	1st Qu.:	0.000000
Median :	4.00	Median :	4.1507	Median :	11.051	Median :	0.000000
Max. :	4.00	Max. :	4.1507	Max. :	11.051	Max. :	0.000000

I		J		K		L	
Min. :	0.0	Min. :	0.000	Min. :	0.000	Min. :	0.000
1st Qu.:	22.0	1st Qu.:	0.000	1st Qu.:	0.775	1st Qu.:	3.045
Median :	121.0	Median :	1.000	Median :	3.817	Median :	11.001
Max. :	121.0	Max. :	1.000	Max. :	9.777	Max. :	11.001

X.1		X.2		X.3		X.4	
Min. :	0.0	Min. :	0.000	Min. :	0.000	Min. :	0.000
1st Qu.:	22.0	1st Qu.:	0.000	1st Qu.:	0.775	1st Qu.:	3.045
Median :	121.0	Median :	1.000	Median :	3.817	Median :	11.001
Max. :	121.0	Max. :	1.000	Max. :	9.777	Max. :	11.001

X.5		X.6		X.7		X.8	
Min. :	0.0000	Min. :	0.0	Min. :	0.00	Min. :	0.000
1st Qu.:	0.0000	1st Qu.:	21.0	1st Qu.:	1.825	1st Qu.:	0.000
Median :	0.0000	Median :	128.0	Median :	9.777	Median :	0.000
Max. :	0.0000	Max. :	-107.0	Max. :	116.0	Max. :	0.000

X.9		X.10		X.11		X.12	
Min. :	0.0000	Min. :	0.0	Min. :	0.00	Min. :	0.000
1st Qu.:	0.0000	1st Qu.:	-369.0	1st Qu.:	19.0	1st Qu.:	0.000
Median :	0.0000	Median :	128.0	Median :	3.00	Median :	0.000
Max. :	0.0000	Max. :	-107.0	Max. :	116.0	Max. :	0.000

X.13		X.14		X.15		X.16	
Min. :	0.0000	Min. :	0.0000	Min. :	0.000000	Min. :	-138.0000
1st Qu.:	0.0000	1st Qu.:	0.0000	1st Qu.:	0.000000	1st Qu.:	0.0000
Median :	0.0000	Median :	0.1087	Median :	0.000000	Median :	0.0000
Max. :	0.0000	Max. :	0.04317	Max. :	0.000000	Max. :	0.0000

X.17		X.18		X.19		X.20	
Min. :	0.0000	Min. :	0.0000	Min. :	0.000000	Min. :	0.0000
1st Qu.:	0.0000	1st Qu.:	0.0000	1st Qu.:	0.000000	1st Qu.:	0.0000
Median :	0.0000	Median :	0.3945	Median :	0.000000	Median :	0.0000
Max. :	0.0000	Max. :	0.04317	Max. :	0.000000	Max. :	0.0000

X.21		X.22		X.23		X.24	
Min. :	0.000	Min. :	0.000000	Min. :	0.000000	Min. :	0.000000
1st Qu.:	0.000	1st Qu.:	0.000000	1st Qu.:	0.000000	1st Qu.:	0.000000
Median :	2.000	Median :	0.000000	Median :	0.03763	Median :	0.000000
Max. :	2.000	Max. :	0.000000	Max. :	0.09264	Max. :	0.000000

X.25		X.26		X.27		X.28	
Min. :	0.000	Min. :	0.000000	Min. :	0.000000	Min. :	0.000000
1st Qu.:	0.000	1st Qu.:	0.000000	1st Qu.:	0.000000	1st Qu.:	0.000000
Median :	2.000	Median :	0.000000	Median :	0.03763	Median :	0.000000
Max. :	2.000	Max. :	0.000000	Max. :	0.09264	Max. :	0.000000

X.29		X.30		X.31		X.32	
Min. :	0.000	Min. :	0.000000	Min. :	0.000000	Min. :	0.000000
1st Qu.:	0.000	1st Qu.:	0.000000	1st Qu.:	0.000000	1st Qu.:	0.000000
Median :	2.000	Median :	0.000000	Median :	0.03763	Median :	0.000000
Max. :	2.000	Max. :	0.000000	Max. :	0.09264	Max. :	0.000000

X.33		X.34		X.35		X.36	
Min. :	0.000	Min. :	0.000000	Min. :	0.000000	Min. :	0.000000
1st Qu.:	0.000	1st Qu.:	0.000000	1st Qu.:	0.000000	1st Qu.:	0.000000
Median :	2.000	Median :	0.000000	Median :	0.03763	Median :	0.000000
Max. :	2.000	Max. :	0.000000	Max. :	0.09264	Max. :	0.000000

Min. : 0.0000	Min. : 0.000000	Min. : 0.000	Min. : 0.0000
Min. : -0.05556	Min. : 0.0000	Min. : -20.000	Min. : 0.000
1st Qu.: 0.0000	1st Qu.: 0.000000	1st Qu.: 0.000	1st Qu.: 0.0000
1st Qu.: 0.00000	1st Qu.: 0.0000	1st Qu.: -7.000	1st Qu.: 0.000
Median : 0.3713	Median : 0.000000	Median : 2.000	Median : 0.0000
Median : 0.00000	Median : 0.3277	Median : -2.000	Median : 2.000
X.37	X.38	X.39	X.40
X.41	X.42	X.43	X.44
Min. : -0.500000	Min. : 0.00	Min. : 0.00	Min. : 0.00
Min. : 0.0	Min. : -1256.000	Min. : 0.0000	Min. : 0.0000
1st Qu.: 0.000000	1st Qu.: 0.00	1st Qu.: 0.00	1st Qu.: 0.00
1st Qu.: 0.0	1st Qu.: -1.000	1st Qu.: 0.0000	1st Qu.: 0.0000
Median : 0.000000	Median : 3.00	Median : 0.00	Median : 0.00
Median : 2.0	Median : 0.000	Median : 0.0000	Median : 0.0000
X.45	X.46	X.47	X.48
X.49	X.50	X.51	X.52
Min. : 0.0000	Min. : 0.0000	Min. : -20.00000	Min. : 0.00
Min. : 0	Min. : 0.000000	Min. : 0.0000	Min. : 0.00e+00
1st Qu.: 0.0000	1st Qu.: 0.0000	1st Qu.: 0.00000	1st Qu.: 14.00
1st Qu.: 93	1st Qu.: 0.000000	1st Qu.: 0.0000	1st Qu.: 0.00e+00
Median : 0.0000	Median : 0.0000	Median : 0.00000	Median : 35.00
Median : 1859	Median : 0.000000	Median : 0.0000	Median : 0.00e+00
X.53	X.54	X.55	X.56
X.57	X.58	X.59	X.60
Min. : 0.0000000	Min. : 0.0000	Min. : 0.0000	Min. : 0.0000
Min. : 0.000000	Min. : 0.000000	Min. : 0.000	Min. : 0.00e+00
1st Qu.: 0.0000000	1st Qu.: 0.0000	1st Qu.: 0.0000	1st Qu.: 0.0000
1st Qu.: 0.000000	1st Qu.: 0.000000	1st Qu.: 0.000	1st Qu.: 0.00e+00
Median : 0.0000000	Median : 0.0000	Median : 0.0000	Median : 0.0000
Median : 0.000000	Median : 0.000000	Median : 0.000	Median : 0.00e+00
X.61	X.62	X.63	X.64
X.65	X.66	X.67	X.68
Min. : 0.0000000	Min. : 0.000000	Min. : 0.00e+00	Min. : 0.0000
Min. : 0.000000	Min. : 0.0000	Min. : 0.00e+00	Min. : 0.00e+00
1st Qu.: 0.0000000	1st Qu.: 0.000000	1st Qu.: 0.00e+00	1st Qu.: 0.0000
1st Qu.: 0.000000	1st Qu.: 0.0000	1st Qu.: 0.00e+00	1st Qu.: 0.00e+00
Median : 0.0000000	Median : 0.000000	Median : 0.00e+00	Median : 0.0000
Median : 0.000000	Median : 0.0000	Median : 0.00e+00	Median : 0.00e+00
X.69	X.70	X.71	X.72
X.73	X.74	X.75	X.76
Min. : 0.00000	Min. : 0.000000	Min. : 0.0000000	Min. : 0.00000
: 0.0000000	Min. : 0.00000	Min. : 0.0000000	Min. : 0.00000
: 0.00000	Min. : 0.00000	Min. : 0.0000000	Min. : 0.00000
1st Qu.: 0.00000	1st Qu.: 0.000000	1st Qu.: 0.0000000	1st Qu.: 0.00000
Qu.: 0.0000000	1st Qu.: 0.00000	1st Qu.: 0.0000000	1st Qu.: 0.00000
1st Qu.: 0.00000	1st Qu.: 0.00000	1st Qu.: 0.0000000	1st Qu.: 0.00000
Median : 0.00000	Median : 0.000000	Median : 0.0000000	Median : 0.00000
: 0.0000000	Median : 0.00000	Median : 0.0000000	Median : 0.00000
Median : 0.00000	Median : 0.00000	Median : 0.0000000	Median : 0.00000
X.77	X.78	X.79	X.80
X.81	X.82	X.83	
Min. : 0.0000000	Min. : 0.0000000	Min. : 0.0000000	Min. : 0.00000
: 0.0000000	Min. : 0.00e+00	Min. : 0.0000000	Min. : 0.00000
1st Qu.: 0.0000000	1st Qu.: 0.0000000	1st Qu.: 0.0000000	1st Qu.: 0.00000
Qu.: 0.000000	1st Qu.: 0.00e+00	1st Qu.: 0.0000000	1st Qu.: 0.00000
Median : 0.0000000	Median : 0.0000000	Median : 0.0000000	Median : 0.00000
: 0.0000000	Median : 0.00e+00	Median : 0.0000000	Median : 0.00000
X.84	X.85	X.86	X.87
X.88	X.89	X.90	X.91
Min. : 0.00e+00	Min. : 0.0000000	Min. : 0.0000000	Min. : 0.00000
: 0.000000	Min. : 0.0000	Min. : 0.00000	Min. : 0.00000
: 0.00000	Min. : 0.0000	Min. : 0.00000	Min. : 0.00000
1st Qu.: 0.00e+00	1st Qu.: 0.0000000	1st Qu.: 0.0000000	1st Qu.: 0.00000
Qu.: 0.000000	1st Qu.: 0.0000	1st Qu.: 0.00000	1st Qu.: 0.00000
Qu.: 0.00000	1st Qu.: 0.0000	1st Qu.: 0.00000	1st Qu.: 0.00000
Median : 0.00e+00	Median : 0.0000000	Median : 0.0000000	Median : 0.00000
: 0.000000	Median : 1.0000	Median : 0.00000	Median : 0.00000
: 0.00000	Median : 0.00000	Median : 0.00000	Median : 0.00000

X.92		X.93		X.94		X.95	
X.96		X.97		X.98		X.99	
Min.	:0.000000	Min.	:0.000000	Min.	:0.000000	Min.	:0.000000
Min.	:0.0000000	Min.	:0.000000	Min.	:0.0000000	Min.	
:0.0000000							
1st Qu.:0.000000		1st Qu.:0.000000		1st Qu.:0.000000		1st Qu.:0.000000	
1st Qu.:0.0000000		1st Qu.:0.000000		1st Qu.:0.0000000		1st	
Qu.:0.0000000							
Median :0.000000		Median :0.000000		Median :0.000000		Median :0.000000	
Median :0.0000000		Median :0.000000		Median :0.0000000		Median	
:0.0000000							
X.100		X.101		X.102		X.103	
X.104		X.105		X.106		X.107	
Min.	:0.0000000	Min.	:0.0000	Min.	:0.000000	Min.	:0.0000
Min.	:0.000000	Min.	:0.000000	Min.	:0.000000	Min.	:0.0000
1st Qu.:0.0000000		1st Qu.:0.0000		1st Qu.:0.000000		1st Qu.:0.0000	
1st Qu.:0.000000		1st Qu.:0.000000		1st Qu.:0.000000		1st Qu.:0.0000	
Median :0.0000000		Median :0.0000		Median :0.000000		Median :0.0000	
Median :0.000000		Median :0.000000		Median :0.000000		Median :0.0000	
X.108		X.109		X.110		X.111	
X.112		X.113		X.114		X.115	
Min.	:0.000000	Min.	:0.0000	Min.	:0.000000	Min.	:0.0000000
Min.	:0.00000	Min.	:0.00000000	Min.	:0.0000000	Min.	:0.000000
1st Qu.:0.000000		1st Qu.:0.0000		1st Qu.:0.000000		1st Qu.:0.0000000	
1st Qu.:0.0000		1st Qu.:0.0000000		1st Qu.:0.000000		1st Qu.:0.000000	
Median :0.000000		Median :0.0000		Median :0.000000		Median :0.0000000	
Median :0.0000		Median :0.00000000		Median :0.000000		Median :0.000000	
X.116		X.117		X.118		X.119	
X.120		X.121		X.122		X.123	
Min.	:0.0000000	Min.	:0.00e+00	Min.	:0.000000	Min.	
:0.00e+00	Min.	:0.0000000	Min.	:0.0000	Min.	:0.000000	Min.
:0.000000							
1st Qu.:0.0000000		1st Qu.:0.00e+00		1st Qu.:0.000000		1st	
Qu.:0.00e+00	1st	Qu.:0.0000000	1st	Qu.:0.0000	1st	Qu.:0.000000	1st
Qu.:0.000000							
Median :0.0000000		Median :0.00e+00		Median :0.000000		Median	
:0.00e+00	Median	:0.0000000	Median	:0.0000	Median	:0.000000	Median
:0.000000							
X.124		X.125		X.126		X.127	
X.128		X.129		X.130		X.131	
Min.	:0.0000000	Min.	:0.000000	Min.	:0.0000	Min.	:0.000000
Min.	:0.000000	Min.	:0.0000000	Min.	:0.0000	Min.	:0.0000000
1st Qu.:0.000000		1st Qu.:0.000000		1st Qu.:0.0000		1st Qu.:0.000000	
1st Qu.:0.000000		1st Qu.:0.000000		1st Qu.:0.0000		1st Qu.:0.0000000	
Median :0.0000000		Median :0.000000		Median :0.0000		Median :0.000000	
Median :0.000000		Median :0.0000000		Median :1.0000		Median :0.0000000	
X.132		X.133		X.134		X.135	
X.136		X.137		X.138		X.139	
Min.	:0.0000000	Min.	:0.0000000	Min.	:0.0000000	Min.	
:0.0000000	Min.	:0.00e+00	Min.	:0.0000000	Min.	:0.0000	Min.
:0.000000							
1st Qu.:0.0000000		1st Qu.:0.0000000		1st Qu.:0.0000000		1st	
Qu.:0.0000000	1st	Qu.:0.00e+00	1st	Qu.:0.0000000	1st	Qu.:0.0000	1st
Qu.:0.000000							
Median :0.0000000		Median :0.0000000		Median :0.0000000		Median	
:0.0000000	Median	:0.00e+00	Median	:0.0000000	Median	:0.0000	Median
:0.000000							
X.140		X.141		X.142		X.143	
X.144		X.145		X.146		X.147	
Min.	:0.000000	Min.	:0.000000	Min.	:0.00000000	Min.	:0.00e+00
Min.	:0.0000000	Min.	:0.0000				

1st Qu.:0.00e+00	1st Qu.:0.000000	1st Qu.:0.000000	1st Qu.:0.000000
1st Qu.:0.000000	1st Qu.:0.00e+00	1st Qu.:0.000000	1st Qu.:0.000000
Median :0.00e+00	Median :0.000000	Median :0.0000	Median :0.000000
Median :0.000000	Median :0.00e+00	Median :0.000000	Median :0.000000
X.156	X.157	X.158	X.159
X.160	X.161	X.162	X.163
Min. :0.00e+00	Min. :0.000000	Min. :0.00000000	Min. :0.00e+00
Min. :0.00000000	Min. :0.0000	Min. :0.00000000	Min. :0.00000000
1st Qu.:0.00e+00	1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.00e+00
1st Qu.:0.00000000	1st Qu.:0.0000	1st Qu.:0.00000000	1st Qu.:0.00000000
Median :0.00e+00	Median :0.000000	Median :0.00000000	Median :0.00e+00
Median :0.00000000	Median :0.0000	Median :0.00000000	Median :0.00000000
X.164	X.165	X.166	X.167
X.168	X.169	X.170	X.171
Min. :0.00000000	Min. :0.00000000	Min. :0.00e+00	Min. :0.000000
Min. :0.000000	Min. :0.00000000	Min. :0.000000	Min. :0.000000
1st Qu.:0.00000000	1st Qu.:0.00000000	1st Qu.:0.00e+00	1st Qu.:0.000000
1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.000000	1st Qu.:0.000000
Median :0.00000000	Median :0.00000000	Median :0.00e+00	Median :0.000000
Median :0.000000	Median :0.00000000	Median :0.000000	Median :0.000000
X.172	X.173	X.174	X.175
X.176	X.177	X.178	X.179
Min. :0.00000000	Min. :0.000000	Min. :0.000000	Min. :0.00000000
Min. :0.000000	Min. :0.00e+00	Min. :0.000000	Min. :0.000000
1st Qu.:0.00000000	1st Qu.:0.000000	1st Qu.:0.000000	1st Qu.:0.00000000
1st Qu.:0.000000	1st Qu.:0.00e+00	1st Qu.:0.000000	1st Qu.:0.000000
Median :0.00000000	Median :0.000000	Median :0.000000	Median :0.00000000
Median :0.000000	Median :0.00e+00	Median :0.000000	Median :0.000000
X.180	X.181	X.182	X.183
X.184	X.185	X.186	X.187
Min. :0.000000	Min. :0.000000	Min. :0.000000	Min. :0.000000
Min. :0.000000	Min. :0.00e+00	Min. :0.00000000	Min. :0.00e+00
1st Qu.:0.000000	1st Qu.:0.000000	1st Qu.:0.000000	1st Qu.:0.000000
1st Qu.:0.000000	1st Qu.:0.00e+00	1st Qu.:0.00000000	1st Qu.:0.00e+00
Median :0.000000	Median :0.000000	Median :0.000000	Median :0.000000
Median :0.000000	Median :0.00e+00	Median :0.00000000	Median :0.00e+00
X.188	X.189	X.190	X.191
X.192	X.193	X.194	X.195
Min. :0.00000000	Min. :0.000000	Min. :0.00000000	Min. :0.00000000
Min. :0.000000	Min. :0.00000000	Min. :0.000000	Min. :0.000000
1st Qu.:0.00000000	1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.00000000
1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.000000	1st Qu.:0.000000
Median :0.00000000	Median :0.000000	Median :0.00000000	Median :0.00000000
Median :0.000000	Median :0.00000000	Median :0.000000	Median :0.000000
X.196	X.197	X.198	X.199
X.200	X.201	X.202	X.203
Min. :0.00000000	Min. :0.000000	Min. :0.00000000	Min. :0.00e+00
Min. :0.000000	Min. :0.00000000	Min. :0.00000000	Min. :0.000000
1st Qu.:0.00000000	1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.00e+00
1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.00000000	1st Qu.:0.000000
Median :0.00000000	Median :0.000000	Median :0.00000000	Median :0.00e+00
Median :0.000000	Median :0.00000000	Median :0.00000000	Median :0.000000
X.204	X.205	X.206	X.207
X.208	X.209	X.210	X.211
Min. :0.00000000	Min. :0.00000000	Min. :0.000000	Min. :0.00000000
Min. :0.000000	Min. :0.00000000	Min. :0.00000000	Min. :0.00000000
1st Qu.:0.00000000	1st Qu.:0.00000000	1st Qu.:0.000000	1st Qu.:0.00000000
1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.00000000	1st Qu.:0.00000000
Median :0.00000000	Median :0.00000000	Median :0.000000	Median :0.00000000
Median :0.000000	Median :0.00000000	Median :0.00000000	Median :0.00000000
X.212	X.213	X.214	X.215
X.216	X.217	X.218	X.219
Min. :0.00000000	Min. :0.000000	Min. :0.00000000	Min. :0.000000
Min. :0.00000000	Min. :0.000000	Min. :0.00000000	Min. :0.000000
1st Qu.:0.00000000	1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.000000
1st Qu.:0.000000	1st Qu.:0.000000	1st Qu.:0.00000000	1st Qu.:0.000000
Median :0.00000000	Median :0.000000	Median :0.00000000	Median :0.000000
Median :0.00000000	Median :0.000000	Median :0.00000000	Median :0.000000

```

X.220      X.221      X.222      X.223
X.224      X.225      X.226      X.227
Min.       :0.0000    Min.       :0.000000    Min.       :0.000000    Min.       :0.00e+00
Min.       :0.000000    Min.       :0.000000    Min.       :0.00000000    Min.       :0.000000
1st Qu.:0.0000    1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.00e+00
1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.000000
Median :0.0000    Median :0.000000    Median :0.000000    Median :0.00e+00
Median :0.000000    Median :0.000000    Median :0.00000000    Median :0.000000
X.228      X.229      X.230      X.231
X.232      X.233      X.234      X.235
Min.       :0.000000    Min.       :0.000000    Min.       :0.00e+00    Min.       :0.00000000
Min.       :0.000000    Min.       :0.000000    Min.       :0.000000    Min.       :0.000000
1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.00e+00    1st Qu.:0.00000000
1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.000000
Median :0.000000    Median :0.000000    Median :0.00e+00    Median :0.00000000
Median :0.000000    Median :1.000000    Median :0.000000    Median :0.000000
X.236      X.237      X.238      X.239
X.240      X.241      X.242      X.243
Min.       :0.000000    Min.       :0.00e+00    Min.       :0.000000    Min.       :0.000000
Min.       :0.00000000    Min.       :0.000000    Min.       :0.000000    Min.       :0.00000000
1st Qu.:0.000000    1st Qu.:0.00e+00    1st Qu.:0.000000    1st Qu.:0.000000
1st Qu.:0.00000000    1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.00000000
Median :0.000000    Median :0.00e+00    Median :0.000000    Median :0.000000
Median :0.00000000    Median :0.000000    Median :0.000000    Median :0.00000000
X.244      X.245      X.246      X.247
X.248      X.249      X.250      X.251
Min.       :0.000000    Min.       :0.000000    Min.       :0.000000    Min.       :0.000000
Min.       :0.000000    Min.       :0.00000000    Min.       :0.000000    Min.       :0.000000
1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.:0.000000
1st Qu.:0.000000    1st Qu.:0.00000000    1st Qu.:0.000000    1st Qu.:0.000000
Median :0.000000    Median :0.000000    Median :0.000000    Median :0.000000
Median :0.000000    Median :0.00000000    Median :0.000000    Median :0.000000
X.252      X.253      X.254      X.255
X.256      X.257      X.258      X.259      X.260
Min.       :0.0000    Min.       :0.0000    Min.       :0.0000    Min.       :0.0000    Min.
:0.0000    Min.       :0.0000    Min.       :0.0000    Min.       :0.0000    Min.
:0.0000
1st Qu.:0.0000    1st Qu.:0.0000    1st Qu.:0.0000    1st Qu.:0.0000    1st
Qu.:0.0000    1st Qu.:0.0000    1st Qu.:0.0000    1st Qu.:0.0000    1st
Qu.:0.0000
Median :0.0000    Median :0.0000    Median :0.0000    Median :0.0000
Median :0.0000    Median :0.0000    Median :0.0000    Median :0.0000    Median
:0.0000
X.261      X.262      X.263      X.264
X.265      X.266      X.267      Target
Min.       :0.0000    Min.       :0.000000    Min.       :0.000000    Min.       : 0.0000
Min.       :0    Min.       : 0.000    Min.       : 0.0000    Min.       : 0.000
1st Qu.:0.0000    1st Qu.:0.000000    1st Qu.:0.000000    1st Qu.: 0.0000
1st Qu.:0    1st Qu.: 0.000    1st Qu.: 0.0000    1st Qu.: 0.000
Median :0.0000    Median :0.000000    Median :0.000000    Median : 0.0000
Median :0    Median : 0.000    Median : 0.0000    Median : 0.000
[ reached getOption("max.print") -- omitted 3 rows ]

```

c. Visualize the dataset and make inferences from that?

Ans : same as Que e

d. Perform any 3 hypothesis tests using columns of your choice, make conclusions

According to t-test, p value is 0.0128 which is less than 0.05. hence there is correlation

```

> t.test(blogData_train$A, mu = 40.30467)

One Sample t-test

data:  blogData_train$A
t = -2.4895, df = 52396, p-value = 0.0128

```

```

alternative hypothesis: true mean is not equal to 40.30467
95 percent confidence interval:
 38.76668 40.12165
sample estimates:
mean of x
 39.44417

```

According to chi square test, p value is 1 which is more than 0.05. hence there is no correlation

```

> chisq.test(blogData_train$A, blogData_train$Target)

Pearson's Chi-squared test

data: blogData_train$A and blogData_train$Target
X-squared = 171770, df = 188780, p-value = 1

```

According to Anova test, p value is 0.000 which is less than 0.05. hence there is correlation

```

> results<- aov(A~Target, data = blogData_train)
> summary(results)
              Df    Sum Sq   Mean Sq    F value    Pr(>F)
Target          1  77304518  77304518    16156 <2e-16 ***
Residuals   52395 250708202      4785
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

e. Create a linear regression model to predict the no.of comments in the next 24hrs (relative to basetime)

```

> summary(mod)

Call:
lm(formula = blogData_train$A ~ blogData_train$Target)

Residuals:
    Min       1Q   Median       3Q      Max
-1311.76  -30.97   -21.92    6.08  1089.09

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   32.553104   0.307019   106.0 <2e-16 ***
blogData_train$Target  1.018677   0.008014  127.1 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 69.17 on 52395 degrees of freedom
Multiple R-squared:  0.2357, Adjusted R-squared:  0.2357
F-statistic: 1.616e+04 on 1 and 52395 DF, p-value: < 2.2e-16

```

For 1 unit increase there will be 1.01 unit of increase in no.of comments

These 3 terms, t-statistics (127.1), Std Error of regression coefficient (0.008014) and Probability (of committing Type I Error) is $2e-06$ [$2/1000000 = 0.00000200 = \text{almost } 0$] are associated with a t test which tests following Null Hypothesis:

Ho: The slope of no.of comments is not significant

Ha: The slope of no.of comments is significant

0.008014 = standard error of regression co-efficient

As p value is 0.000 which is less than 0.05, Reject the Ho (and accept Ha) and conclude that "slope is significant"

Multiple R-squared = 0.2357

23.5% of the variance in Average can be explained by no.of comments [Remaining 76.5% is unexplained variance....due to factors outside the model]

```
> pred<- predict(lm(blogData_train$A~blogData_train$Target))
> pred
```

1	2	3	4	5	6	7	15
33.57178	32.55310	32.55310	33.57178	60.05738	32.55310	32.55310	
60.05738	41.72120	41.72120	32.55310	32.55310	34.59046	32.55310	
45.79590							
16	17	18	19	20	21	22	30
47.83326	35.60913	47.83326	45.79590	33.57178	53.94532	33.57178	
53.94532	33.57178	35.60913	35.60913	35.60913	35.60913	34.59046	
34.59046							
31	32	33	34	35	36	37	45
37.64649	37.64649	32.55310	32.55310	53.94532	32.55310	32.55310	
35.60913	32.55310	32.55310	32.55310	34.59046	32.55310	32.55310	
32.55310							
46	47	48	49	50	51	52	60
32.55310	46.81458	44.77723	39.68384	32.55310	32.55310	32.55310	
32.55310	32.55310	32.55310	32.55310	32.55310	32.55310	39.68384	
35.60913							
61	62	63	64	65	66	67	75
33.57178	33.57178	33.57178	34.59046	34.59046	32.55310	32.55310	
37.64649	49.87061	37.64649	32.55310	34.59046	32.55310	33.57178	
32.55310							
76	77	78	79	80	81	82	90
33.57178	40.70252	32.55310	32.55310	69.22547	33.57178	34.59046	
44.77723	41.72120	40.70252	41.72120	40.70252	36.62781	40.70252	
33.57178							
91	92	93	94	95	96	97	105
33.57178	35.60913	35.60913	34.59046	34.59046	33.57178	33.57178	
35.60913	35.60913	34.59046	34.59046	35.60913	35.60913	59.03870	
34.59046							
106	107	108	109	110	111	112	120
33.57178	37.64649	33.57178	35.60913	34.59046	35.60913	57.00135	
57.00135	34.59046	34.59046	35.60913	33.57178	104.87916	32.55310	
32.55310							
121	122	123	124	125	126	127	135
36.62781	37.64649	37.64649	37.64649	36.62781	36.62781	36.62781	
38.66517	38.66517	38.66517	78.39356	78.39356	40.70252	40.70252	
32.55310							
136	137	138	139	140	141	142	150
32.55310	36.62781	36.62781	39.68384	32.55310	32.55310	32.55310	
32.55310	32.55310	32.55310	44.77723	32.55310	38.66517	33.57178	
36.62781							
151	152	153	154	155	156	157	165
33.57178	32.55310	32.55310	38.66517	32.55310	38.66517	34.59046	
34.59046	39.68384	39.68384	37.64649	37.64649	36.62781	36.62781	
33.57178							
166	167	168	169	170	171	172	180
173	174	175	176	177	178	179	

33.57178	33.57178	32.55310	107.93519	32.55310	107.93519	52.92664	
52.92664	37.64649	37.64649	50.88929	50.88929	33.57178	33.57178	
40.70252							
181	182	183	184	185	186	187	
188	189	190	191	192	193	194	195
32.55310	32.55310	34.59046	32.55310	32.55310	38.66517	32.55310	
32.55310	32.55310	34.59046	33.57178	33.57178	32.55310	32.55310	
32.55310							
196	197	198	199	200	201	202	
203	204	205	206	207	208	209	210
33.57178	33.57178	32.55310	32.55310	32.55310	32.55310	32.55310	
32.55310	32.55310	32.55310	32.55310	32.55310	55.98267	40.70252	
37.64649							
211	212	213	214	215	216	217	
218	219	220	221	222	223	224	225
36.62781	36.62781	36.62781	36.62781	32.55310	35.60913	32.55310	
35.60913	36.62781	36.62781	80.43091	80.43091	33.57178	80.43091	
33.57178							
226	227	228	229	230	231	232	
233	234	235	236	237	238	239	240
36.62781	33.57178	33.57178	36.62781	112.00989	112.00989	112.00989	
33.57178	32.55310	32.55310	32.55310	32.55310	41.72120	41.72120	
41.72120							
241	242	243	244	245	246	247	
248	249	250	251	252	253	254	255
32.55310	32.55310	32.55310	35.60913	35.60913	35.60913	33.57178	
33.57178	33.57178	33.57178	35.60913	32.55310	32.55310	35.60913	
35.60913							
256	257	258	259	260	261	262	
263	264	265	266	267	268	269	270
33.57178	33.57178	33.57178	32.55310	32.55310	32.55310	32.55310	
32.55310	32.55310	39.68384	39.68384	39.68384	36.62781	36.62781	
36.62781							
271	272	273	274	275	276	277	
278	279	280	281	282	283	284	285
39.68384	39.68384	39.68384	35.60913	35.60913	33.57178	33.57178	
44.77723	32.55310	32.55310	44.77723	38.66517	38.66517	36.62781	
36.62781							
286	287	288	289	290	291	292	
293	294	295	296	297	298	299	300
38.66517	38.66517	38.66517	33.57178	33.57178	33.57178	35.60913	
35.60913	35.60913	36.62781	36.62781	35.60913	35.60913	32.55310	
32.55310							
301	302	303	304	305	306	307	
308	309	310	311	312	313	314	315
34.59046	34.59046	34.59046	43.75855	43.75855	35.60913	35.60913	
35.60913	34.59046	34.59046	32.55310	32.55310	32.55310	34.59046	
34.59046							
316	317	318	319	320	321	322	
323	324	325	326	327	328	329	330
35.60913	35.60913	33.57178	33.57178	32.55310	32.55310	41.72120	
41.72120	41.72120	39.68384	39.68384	41.72120	34.59046	69.22547	
69.22547							
331	332	333	334	335	336	337	
338	339	340	341	342	343	344	345
34.59046	34.59046	34.59046	42.73987	40.70252	34.59046	40.70252	
32.55310	34.59046	32.55310	34.59046	34.59046	41.72120	41.72120	
40.70252							
346	347	348	349	350	351	352	
353	354	355	356	357	358	359	360
39.68384	39.68384	40.70252	32.55310	37.64649	37.64649	32.55310	
32.55310	33.57178	33.57178	32.55310	32.55310	33.57178	32.55310	
35.60913							
361	362	363	364	365	366	367	
368	369	370	371	372	373	374	375
32.55310	35.60913	33.57178	88.58033	88.58033	46.81458	46.81458	
32.55310	32.55310	40.70252	36.62781	33.57178	43.75855	34.59046	
32.55310							

383	376	384	377	385	378	386	379	387	380	388	381	389	382	390
36.62781	34.59046	33.57178	34.59046	33.57178	34.59046	34.59046	34.59046	32.55310	32.55310	32.55310	32.55310	51.90796		
32.55310	32.55310	42.73987	42.73987	42.73987	42.73987	42.73987	42.73987	42.73987	42.73987	42.73987	42.73987	51.90796		
37.64649														
391	392	393	394	395	396	397	398	399	400	401	402	403	404	405
51.90796	47.83326	37.64649	47.83326	47.83326	47.83326	35.60913	36.62781	38.66517	35.60913	35.60913	35.60913	35.60913	36.62781	
36.62781	38.66517	35.60913	35.60913	35.60913	35.60913	35.60913	35.60913	35.60913	35.60913	35.60913	35.60913	35.60913	36.62781	
35.60913														
406	407	408	409	410	411	412	413	414	415	416	417	418	419	420
33.57178	40.70252	40.70252	42.73987	33.57178	32.55310	36.62781	32.55310	42.73987	33.57178	32.55310	32.55310	32.55310	32.55310	
32.55310	42.73987	33.57178	32.55310	36.62781	32.55310	36.62781	32.55310	42.73987	33.57178	32.55310	32.55310	32.55310	32.55310	
32.55310														
421	422	423	424	425	426	427	428	429	430	431	432	433	434	435
37.64649	35.60913	42.73987	32.55310	32.55310	32.55310	32.55310	32.55310	35.60913	42.73987	32.55310	32.55310	32.55310	32.55310	
32.55310	32.55310	34.59046	34.59046	40.70252	51.90796	40.70252	51.90796	32.55310	34.59046	32.55310	32.55310	32.55310	32.55310	
51.90796														
436	437	438	439	440	441	442	443	444	445	446	447	448	449	450
33.57178	36.62781	33.57178	36.62781	41.72120	41.72120	32.55310	32.55310	38.66517	33.57178	36.62781	41.72120	41.72120	32.55310	
32.55310	38.66517	38.66517	33.57178	33.57178	33.57178	33.57178	33.57178	38.66517	33.57178	36.62781	41.72120	41.72120	32.55310	
34.59046														
451	452	453	454	455	456	457	458	459	460	461	462	463	464	465
34.59046	35.60913	35.60913	92.65503	92.65503	32.55310	32.55310	50.88929	50.88929	36.62781	32.55310	32.55310	36.62781	39.68384	
50.88929	50.88929	36.62781	32.55310	32.55310	36.62781	39.68384	38.66517	50.88929	36.62781	32.55310	32.55310	36.62781	39.68384	
38.66517														
466	467	468	469	470	471	472	473	474	475	476	477	478	479	480
38.66517	36.62781	39.68384	32.55310	32.55310	36.62781	37.64649	35.60913	37.64649	35.60913	38.66517	38.66517	158.86902	158.86902	
35.60913	37.64649	35.60913	38.66517	38.66517	36.62781	37.64649	35.60913	37.64649	35.60913	38.66517	38.66517	158.86902	158.86902	
158.86902														
481	482	483	484	485	486	487	488	489	490	491	492	493	494	495
158.86902	88.58033	88.58033	88.58033	88.58033	53.94532	53.94532	53.94532	53.94532	41.72120	41.72120	33.57178	33.57178	33.57178	
53.94532	53.94532	41.72120	41.72120	41.72120	53.94532	53.94532	53.94532	53.94532	41.72120	41.72120	33.57178	33.57178	33.57178	
33.57178														
496	497	498	499	500	501	502	503	504	505	506	507	508	509	510
64.13208	64.13208	39.68384	39.68384	35.60913	35.60913	34.59046	34.59046	32.55310	32.55310	40.70252	40.70252	36.62781	36.62781	
34.59046	64.13208	39.68384	39.68384	35.60913	35.60913	34.59046	34.59046	32.55310	32.55310	40.70252	40.70252	36.62781	36.62781	
36.62781														
511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
41.72120	41.72120	41.72120	35.60913	35.60913	35.60913	32.55310	32.55310	41.72120	41.72120	40.70252	40.70252	35.60913	32.55310	
32.55310	41.72120	41.72120	35.60913	35.60913	35.60913	32.55310	32.55310	41.72120	41.72120	40.70252	40.70252	35.60913	32.55310	
32.55310														
526	527	528	529	530	531	532	533	534	535	536	537	538	539	540
32.55310	32.55310	32.55310	32.55310	32.55310	32.55310	41.72120	64.13208	42.73987	45.79590	45.79590	45.79590	44.77723	44.77723	
64.13208	32.55310	32.55310	32.55310	32.55310	32.55310	41.72120	64.13208	42.73987	45.79590	45.79590	45.79590	44.77723	44.77723	
44.77723														
541	542	543	544	545	546	547	548	549	550	551	552	553	554	555
37.64649	37.64649	37.64649	35.60913	35.60913	72.28150	35.60913	35.60913	72.28150	72.28150	73.30018	38.66517	38.66517	73.30018	
35.60913	37.64649	37.64649	35.60913	35.60913	72.28150	35.60913	35.60913	72.28150	72.28150	73.30018	38.66517	38.66517	73.30018	
73.30018														
556	557	558	559	560	561	562	563	564	565	566	567	568	569	570
48.85193	48.85193	48.85193	50.88929	37.64649	33.57178	32.55310	37.64649	48.85193	48.85193	50.88929	37.64649	33.57178	32.55310	
37.64649	48.85193	48.85193	50.88929	37.64649	33.57178	32.55310	37.64649	48.85193	48.85193	50.88929	37.64649	33.57178	32.55310	
33.57178														
571	572	573	574	575	576	577	578	579	580	581	582	583	584	585
578	579	580	581	582	583	584	585							

33.57178	34.59046	34.59046	42.73987	42.73987	42.73987	38.66517	
38.66517	38.66517	33.57178	33.57178	33.57178	54.96399	54.96399	
33.57178							
586	587	588	589	590	591	592	
593	594	595	596	597	598	599	600
33.57178	34.59046	34.59046	33.57178	33.57178	38.66517	32.55310	
32.55310	38.66517	38.66517	38.66517	34.59046	32.55310	32.55310	
34.59046							
601	602	603	604	605	606	607	
608	609	610	611	612	613	614	615
34.59046	34.59046	34.59046	34.59046	34.59046	34.59046	32.55310	
32.55310	46.81458	33.57178	46.81458	33.57178	37.64649	32.55310	
37.64649							
616	617	618	619	620	621	622	
623	624	625	626	627	628	629	630
32.55310	34.59046	34.59046	50.88929	50.88929	38.66517	39.68384	
45.79590	45.79590	45.79590	39.68384	38.66517	33.57178	35.60913	
32.55310							
631	632	633	634	635	636	637	
638	639	640	641	642	643	644	645
32.55310	32.55310	35.60913	33.57178	32.55310	32.55310	32.55310	
32.55310	32.55310	39.68384	39.68384	33.57178	33.57178	33.57178	
33.57178							
646	647	648	649	650	651	652	
653	654	655	656	657	658	659	660
33.57178	33.57178	32.55310	32.55310	32.55310	32.55310	71.26282	
35.60913	36.62781	32.55310	53.94532	32.55310	43.75855	57.00135	
42.73987							
661	662	663	664	665	666	667	
668	669	670	671	672	673	674	675
33.57178	33.57178	32.55310	32.55310	32.55310	32.55310	54.96399	
54.96399	32.55310	32.55310	35.60913	35.60913	58.02002	74.31885	
46.81458							
676	677	678	679	680	681	682	
683	684	685	686	687	688	689	690
34.59046	34.59046	32.55310	32.55310	33.57178	32.55310	33.57178	
32.55310	32.55310	32.55310	32.55310	32.55310	34.59046	34.59046	
34.59046							
691	692	693	694	695	696	697	
698	699	700	701	702	703	704	705
32.55310	32.55310	32.55310	34.59046	34.59046	34.59046	32.55310	
34.59046	34.59046	32.55310	34.59046	65.15076	65.15076	32.55310	
65.15076							
706	707	708	709	710	711	712	
713	714	715	716	717	718	719	720
42.73987	42.73987	42.73987	34.59046	34.59046	34.59046	38.66517	
38.66517	38.66517	39.68384	34.59046	34.59046	39.68384	39.68384	
32.55310							
721	722	723	724	725	726	727	
728	729	730	731	732	733	734	735
32.55310	32.55310	32.55310	33.57178	33.57178	33.57178	34.59046	
34.59046	32.55310	34.59046	32.55310	32.55310	33.57178	33.57178	
32.55310							
736	737	738	739	740	741	742	
743	744	745	746	747	748	749	750
33.57178	32.55310	32.55310	34.59046	34.59046	34.59046	36.62781	
36.62781	36.62781	36.62781	36.62781	36.62781	32.55310	32.55310	
32.55310							
751	752	753	754	755	756	757	
758	759	760	761	762	763	764	765
33.57178	33.57178	33.57178	34.59046	34.59046	34.59046	32.55310	
32.55310	32.55310	34.59046	34.59046	34.59046	39.68384	39.68384	
39.68384							
766	767	768	769	770	771	772	
773	774	775	776	777	778	779	780
32.55310	32.55310	32.55310	38.66517	38.66517	38.66517	32.55310	
32.55310	32.55310	32.55310	32.55310	36.62781	36.62781	36.62781	
32.55310							

788	781	782	783	784	785	786	787	795
32.55310	44.77723	44.77723	44.77723	32.55310	32.55310	32.55310	33.57178	
57.00135	57.00135	57.00135	33.57178	33.57178	32.55310	32.55310	32.55310	
32.55310								
796	797	798	799	800	801	802		
803	804	805	806	807	808	809	810	
38.66517	38.66517	38.66517	35.60913	35.60913	35.60913	32.55310		
32.55310	32.55310	32.55310	32.55310	32.55310	35.60913	35.60913		
32.55310								
811	812	813	814	815	816	817		
818	819	820	821	822	823	824	825	
35.60913	32.55310	32.55310	33.57178	33.57178	33.57178	33.57178		
33.57178	33.57178	33.57178	33.57178	33.57178	33.57178	33.57178		
33.57178								
826	827	828	829	830	831	832		
833	834	835	836	837	838	839	840	
34.59046	34.59046	34.59046	32.55310	32.55310	32.55310	32.55310		
32.55310	32.55310	32.55310	32.55310	32.55310	32.55310	32.55310		
32.55310								
841	842	843	844	845	846	847		
848	849	850	851	852	853	854	855	
32.55310	32.55310	32.55310	33.57178	33.57178	32.55310	32.55310		
32.55310	32.55310	32.55310	32.55310	33.57178	42.73987	42.73987		
42.73987								
856	857	858	859	860	861	862		
863	864	865	866	867	868	869	870	
37.64649	33.57178	33.57178	37.64649	33.57178	37.64649	35.60913		
35.60913	35.60913	33.57178	33.57178	33.57178	32.55310	32.55310		
32.55310								
871	872	873	874	875	876	877		
878	879	880	881	882	883	884	885	
32.55310	32.55310	32.55310	33.57178	33.57178	33.57178	32.55310		
32.55310	39.68384	32.55310	39.68384	32.55310	32.55310	35.60913		
32.55310								
886	887	888	889	890	891	892		
893	894	895	896	897	898	899	900	
35.60913	32.55310	32.55310	33.57178	32.55310	33.57178	33.57178		
33.57178	35.60913	33.57178	35.60913	32.55310	32.55310	32.55310		
32.55310								
901	902	903	904	905	906	907		
908	909	910	911	912	913	914	915	
32.55310	32.55310	34.59046	32.55310	32.55310	34.59046	32.55310		
34.59046	32.55310	33.57178	34.59046	34.59046	33.57178	33.57178		
34.59046								
916	917	918	919	920	921	922		
923	924	925	926	927	928	929	930	
32.55310	33.57178	33.57178	32.55310	32.55310	33.57178	34.59046		
33.57178	33.57178	34.59046	33.57178	34.59046	33.57178	33.57178		
33.57178								
931	932	933	934	935	936	937		
938	939	940	941	942	943	944	945	
32.55310	44.77723	44.77723	44.77723	32.55310	32.55310	32.55310		
32.55310	32.55310	36.62781	32.55310	32.55310	36.62781	32.55310		
36.62781								
946	947	948	949	950	951	952		
953	954	955	956	957	958	959	960	
32.55310	32.55310	32.55310	35.60913	35.60913	35.60913	33.57178		
33.57178	33.57178	33.57178	55.98267	55.98267	33.57178	55.98267		
33.57178								
961	962	963	964	965	966	967		
968	969	970	971	972	973	974	975	
55.98267	34.59046	46.81458	46.81458	34.59046	46.81458	34.59046		
46.81458	37.64649	37.64649	32.55310	37.64649	32.55310	32.55310		
37.64649								
976	977	978	979	980	981	982		
983	984	985	986	987	988	989	990	

```
32.55310 32.55310 32.55310 32.55310 32.55310 32.55310 32.55310
32.55310 32.55310 34.59046 32.55310 32.55310 34.59046 32.55310
32.55310
991      992      993      994      995      996      997
998      999      1000
34.59046 32.55310 32.55310 32.55310 32.55310 32.55310 32.55310
32.55310 32.55310 32.55310
[ reached getOption("max.print") -- omitted 51397 entries ]
```