

```
In [51]: import pandas as pd
from pandas_profiling import ProfileReport
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
sns.set(color_codes=True)
%matplotlib inline
```

```
In [52]: data = pd.read_csv('turkey_quake.csv')
```

```
In [53]: data.head(2)
```

```
Out[53]:
```

	time	latitude	longitude	depth	mag	magType	nst	gap	dmin	rms	...	
0	2023-02-07T05:26:31.390Z	36.1206	36.3448	18.551	4.4		mb	50.0	128.0	1.259	0.48	...

1	2023-02-07T04:16:42.785Z	37.8886	37.3491	10.000	4.3		mb	20.0	87.0	0.724	0.36	...
---	--------------------------	---------	---------	--------	-----	--	----	------	------	-------	------	-----

2 rows × 22 columns

```
In [54]: pd.isnull(data).sum().sum()
```

```
Out[54]: 22
```

```
In [55]: data.dropna(inplace=True)
```

```
In [56]: data.head(2)
```

```
Out[56]:
```

	time	latitude	longitude	depth	mag	magType	nst	gap	dmin	rms	...
--	------	----------	-----------	-------	-----	---------	-----	-----	------	-----	-----

0	2023-02-07T05:26:31.390Z	36.1206	36.3448	18.551	4.4		mb	50.0	128.0	1.259	0.48	...
---	--------------------------	---------	---------	--------	-----	--	----	------	-------	-------	------	-----

1	2023-02-07T04:16:42.785Z	37.8886	37.3491	10.000	4.3		mb	20.0	87.0	0.724	0.36	...
---	--------------------------	---------	---------	--------	-----	--	----	------	------	-------	------	-----

2 rows × 22 columns

```
In [57]: pd.isnull(data).sum().sum()
```

```
Out[57]: 0
```

```
In [58]: data.shape
```

```
Out[58]: (98, 22)
```

```
In [9]: ProfileReport(data)
```

```
Summarize dataset: 0% | 0/5 [00:00<?, ?it/s]  
Generate report structure: 0% | 0/1 [00:00<?, ?it/s]  
Render HTML: 0% | 0/1 [00:00<?, ?it/s]
```

Overview

Dataset statistics

Number of variables	22
Number of observations	98
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	19.7 KiB
Average record size in memory	205.5 B

Variable types

Categorical	10
Numeric	12

Alerts

type has constant value "earthquake"	Constant
time has a high cardinality: 98 distinct values	High cardinality
id has a high cardinality: 98 distinct values	High cardinality

Out[9]:

In [59]: `from autoviz.AutoViz_Class import AutoViz_Class`In [60]: `AV=AutoViz_Class()`In [61]: `data = AV.AutoViz('turkey_quake.csv')`

```
Shape of your Data Set loaded: (104, 22)
#####
## CLASSIFYING VARIABLES #####
##
##### CLASSIFYING VARIABLES #####
##
Classifying variables in data set...
Data cleaning improvement suggestions. Complete them before proceeding to ML modeling.
```

	Nuniques	dtype	Nulls	Nullpercent	NuniquePercent	Value counts	Data cleaning improvement suggestions
						Min	
time	104	object	0	0.000000	100.000000	1	combine rare categories, possible ID column: drop
latitude	104	float64	0	0.000000	100.000000	0	skewed: cap or drop outliers
updated	104	object	0	0.000000	100.000000	1	combine rare categories, possible ID column: drop
id	104	object	0	0.000000	100.000000	1	combine rare categories, possible ID column: drop
longitude	104	float64	0	0.000000	100.000000	0	skewed: cap or drop outliers
depthError	97	float64	0	0.000000	93.269231	0	skewed: cap or drop outliers
horizontalError	96	float64	3	2.884615	92.307692	0	fill missing
dmin	96	float64	5	4.807692	92.307692	0	fill missing, highly skewed: drop outliers or do box-cox transform
place	94	object	0	0.000000	90.384615	1	combine rare categories
gap	74	float64	3	2.884615	71.153846	0	fill missing, skewed: cap or drop outliers
magError	73	float64	4	3.846154	70.192308	0	fill missing, highly skewed: drop outliers or do box-cox transform
rms	68	float64	0	0.000000	65.384615	0	
magNst	65	float64	4	3.846154	62.500000	0	fill missing, skewed: cap or drop outliers
nst	64	float64	3	2.884615	61.538462	0	fill missing
depth	44	float64	0	0.000000	42.307692	0	skewed: cap or drop outliers
mag	28	float64	0	0.000000	26.923077	0	
magSource	7	object	0	0.000000	6.730769	1	
magType	6	object	0	0.000000	5.769231	1	
net	6	object	0	0.000000	5.769231	1	

	Nuniques	dtype	Nulls	Nullpercent	NuniquePercent	Value counts Min	Data cleaning improvement suggestions
locationSource	6	object	0	0.000000	5.769231	1	
status	2	object	0	0.000000	1.923077	5	
type	1	object	0	0.000000	0.961538	104	invariant values: drop

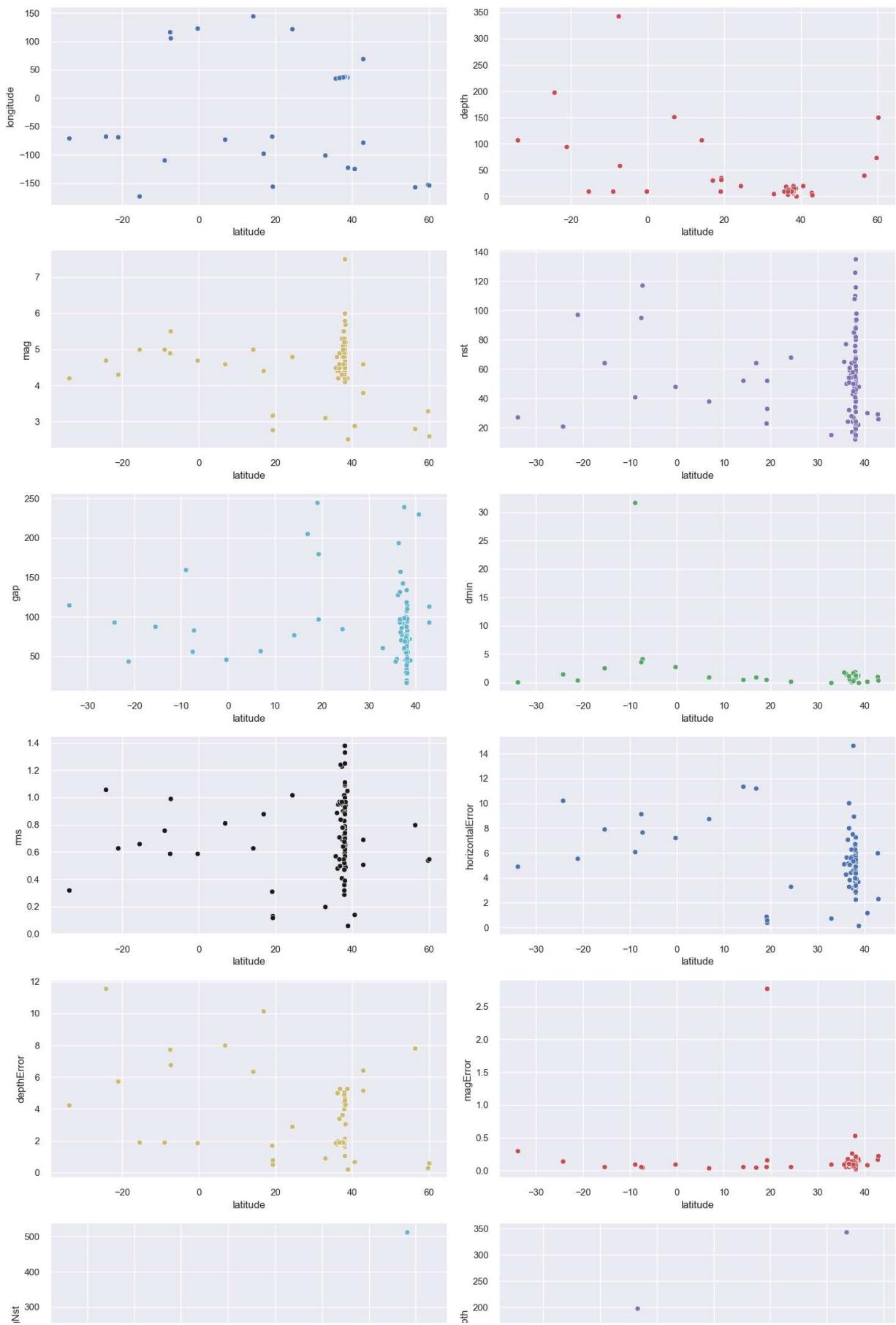
22 Predictors classified...

4 variables removed since they were ID or low-information variables

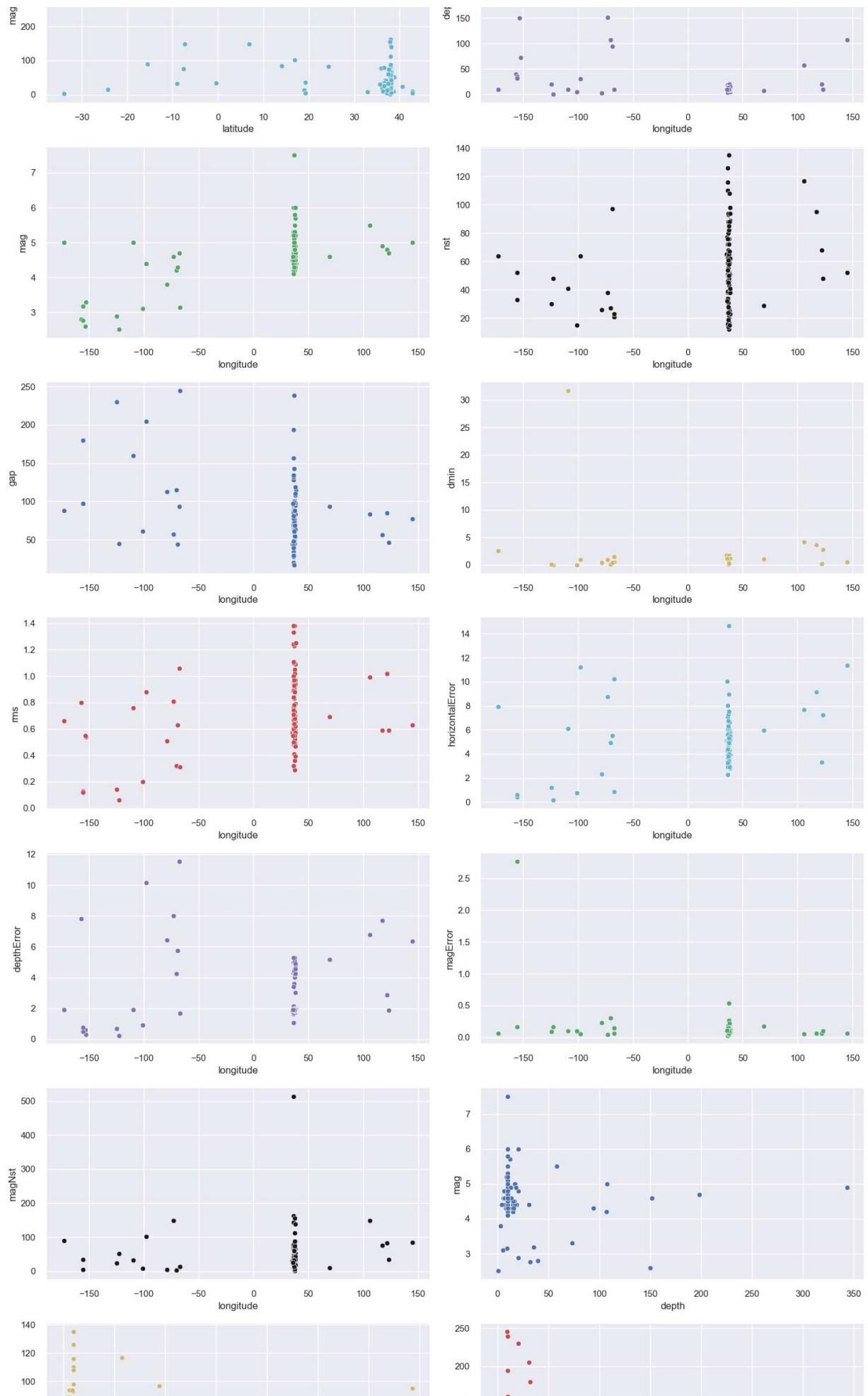
List of variables removed: ['time', 'id', 'updated', 'type']

Number of All Scatter Plots = 78

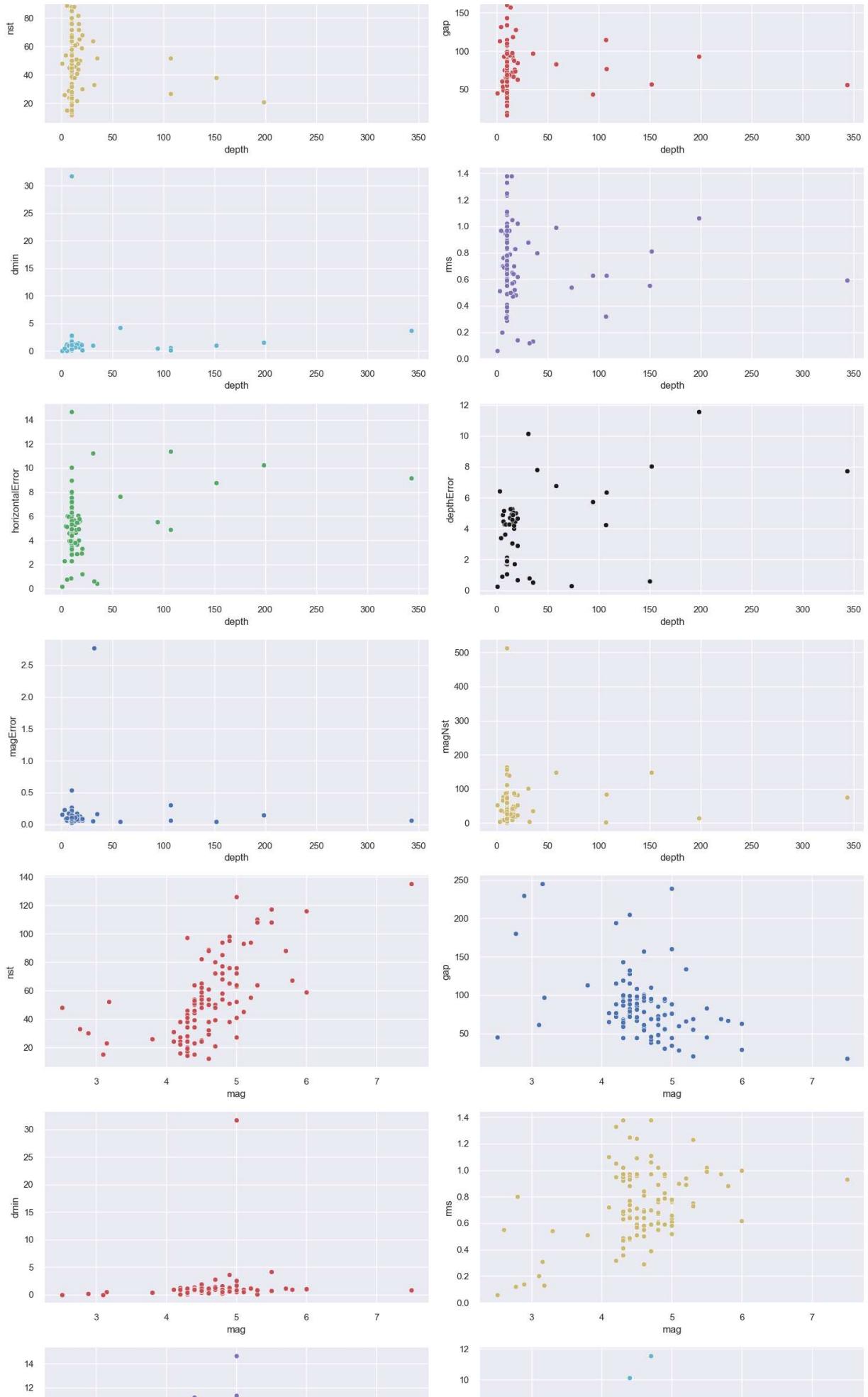
Pair-wise Scatter Plot of all Continuous Variables



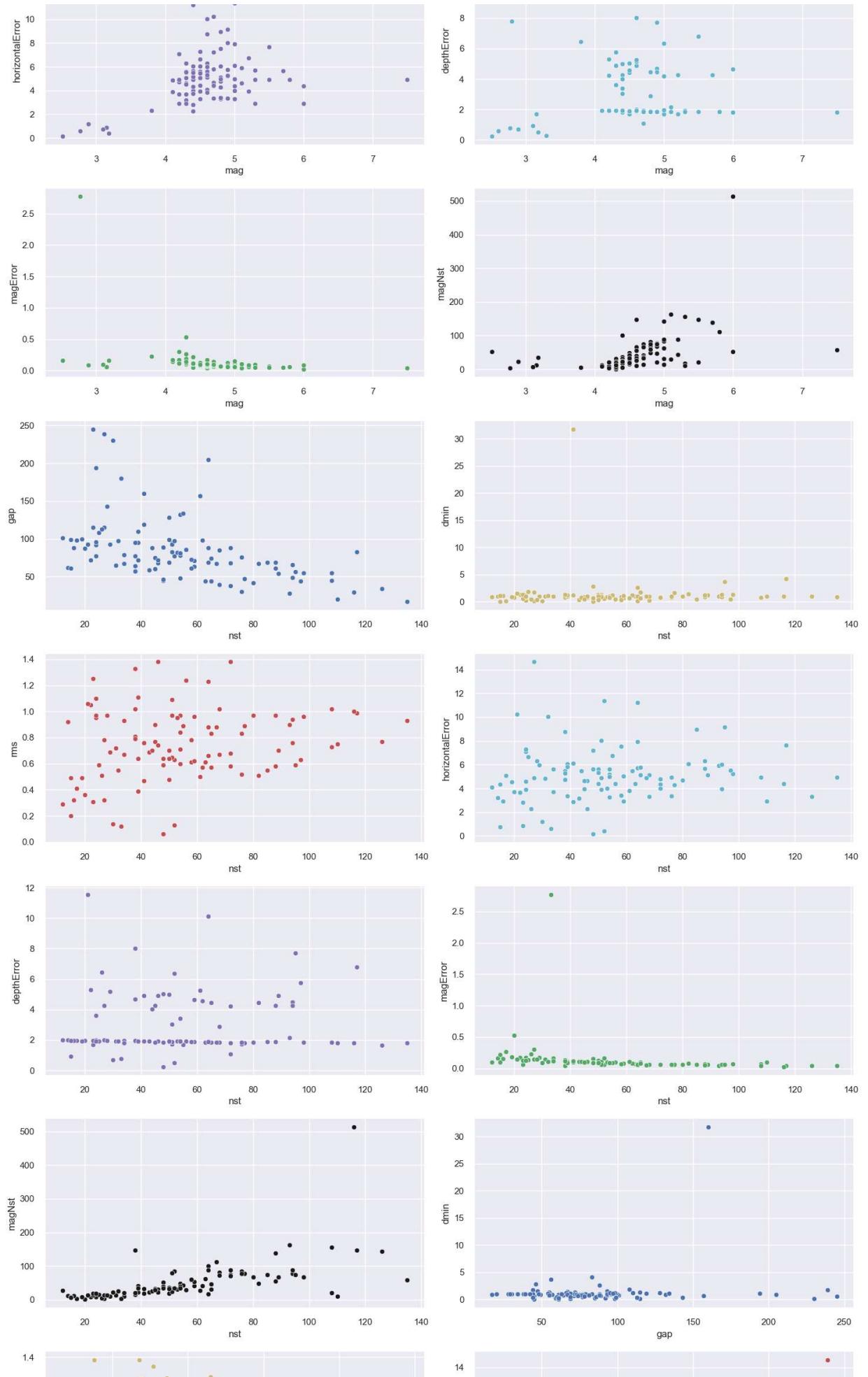
Turkey&Syria earthquake analysis



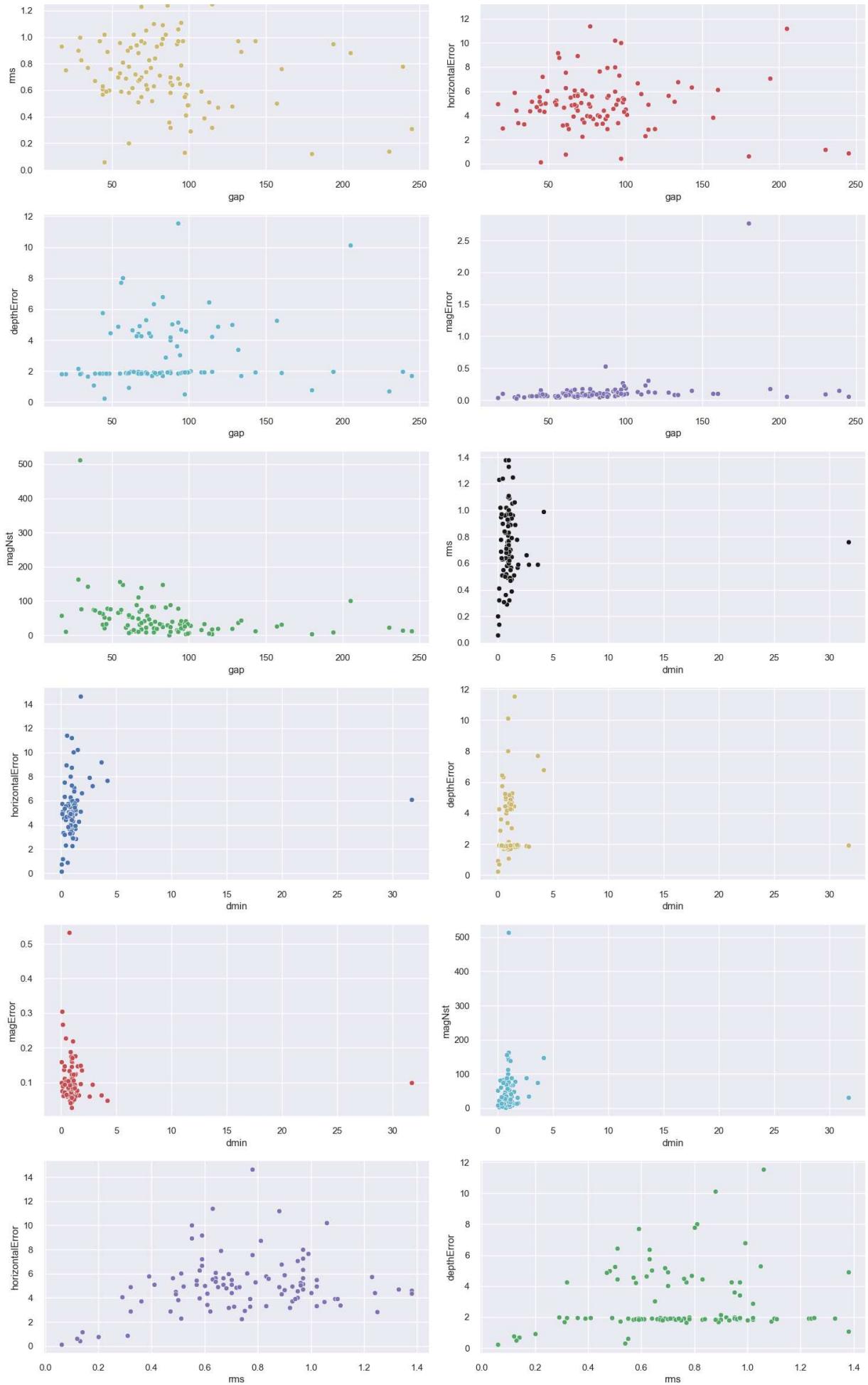
Turkey&Syria earthquake analysis

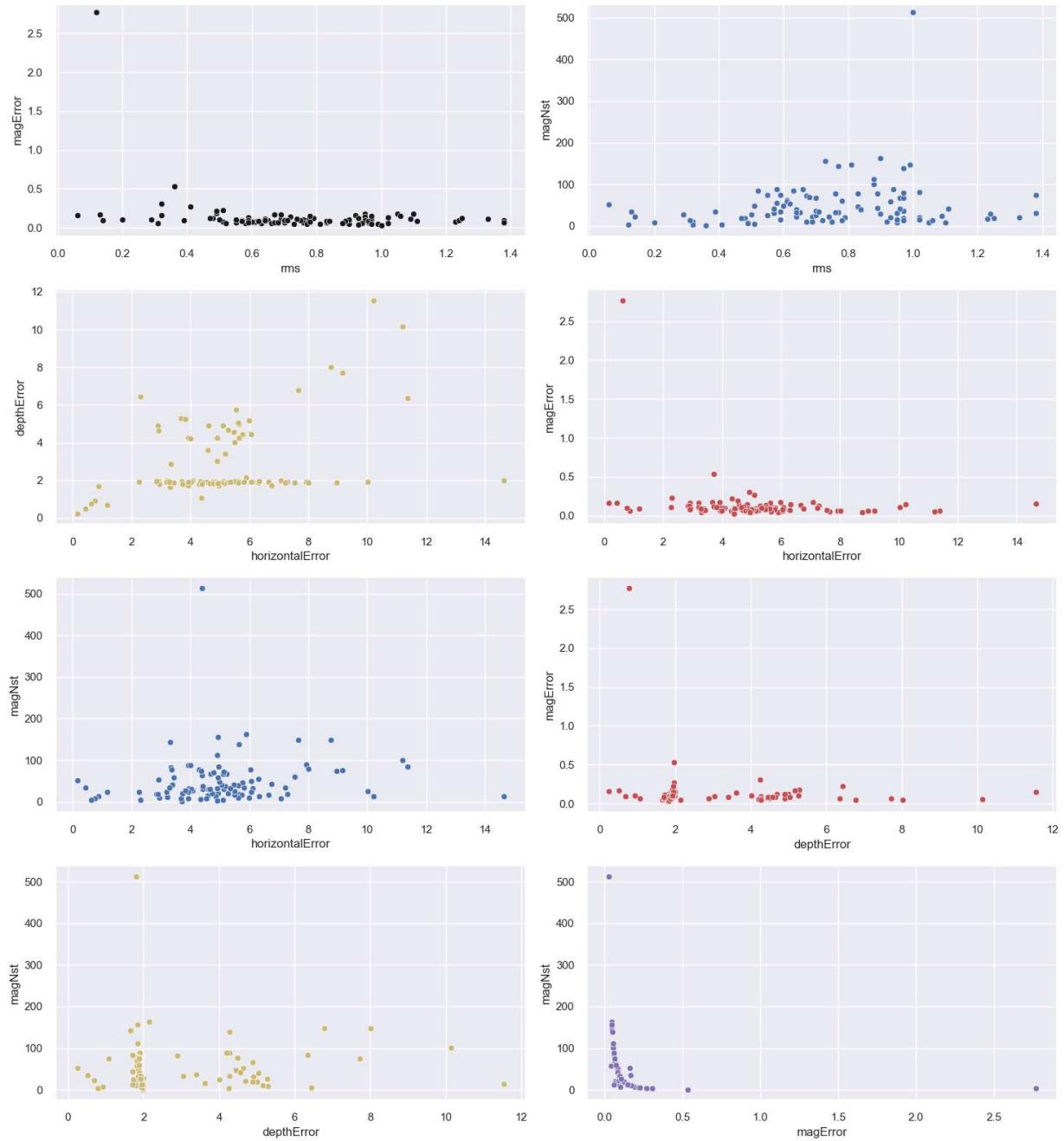


Turkey&Syria earthquake analysis

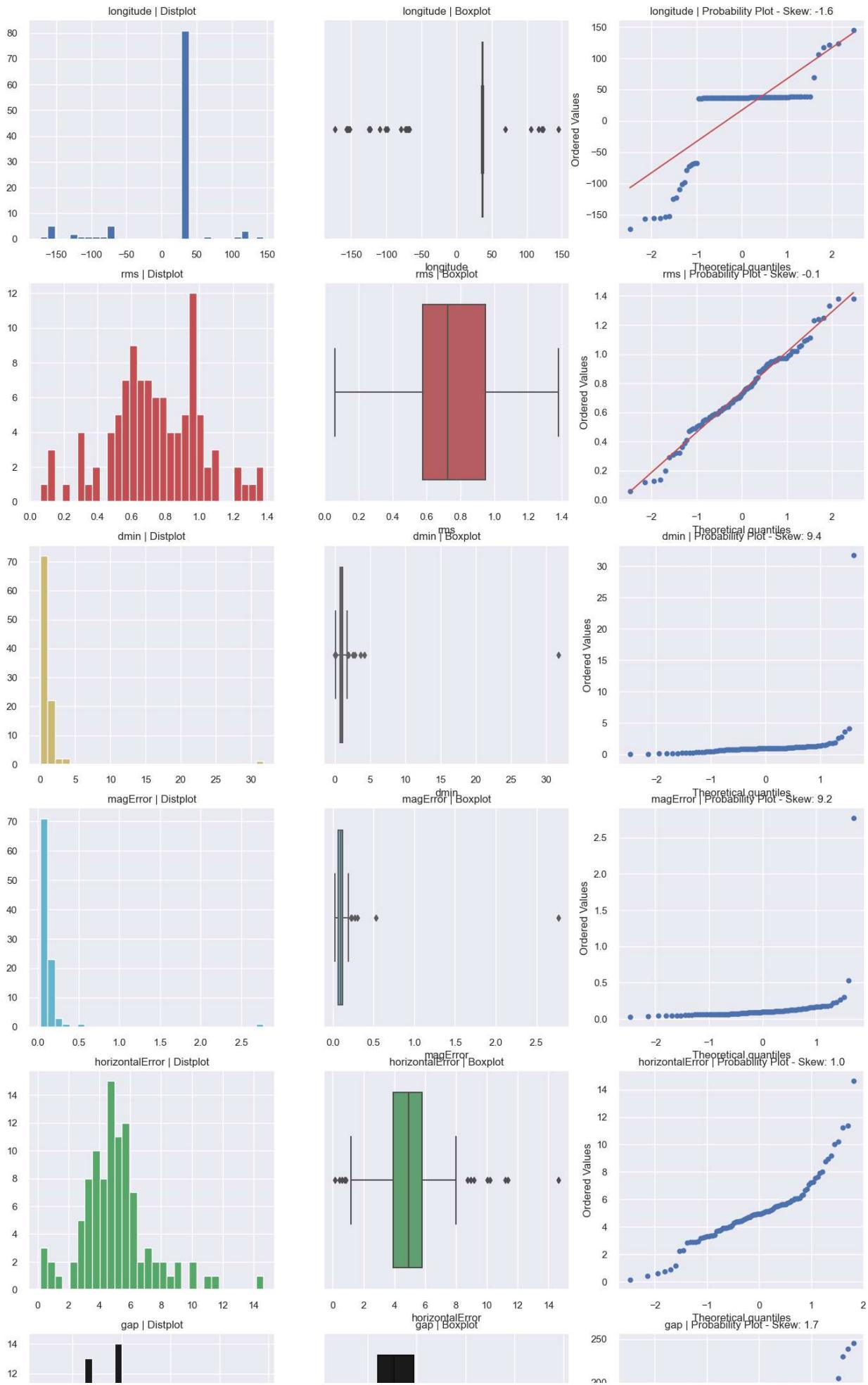


Turkey&Syria earthquake analysis

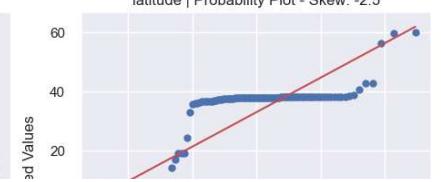
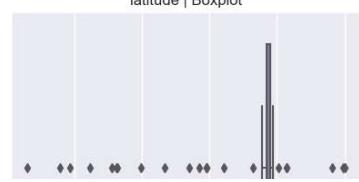
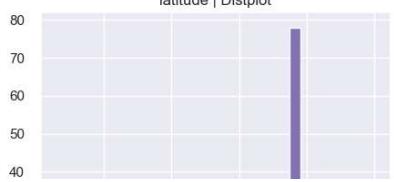
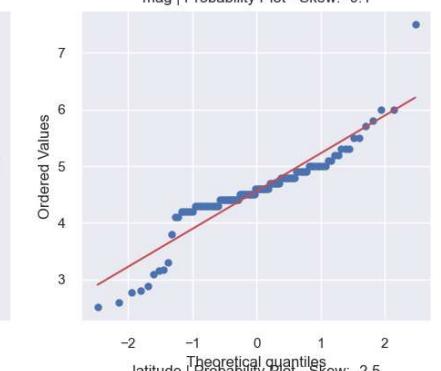
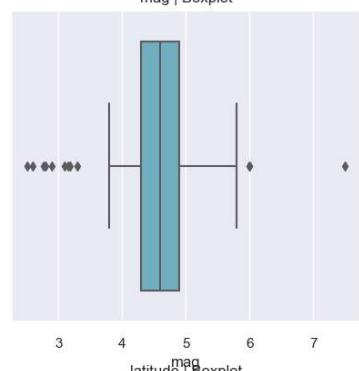
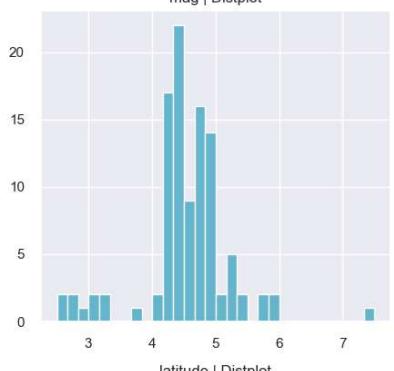
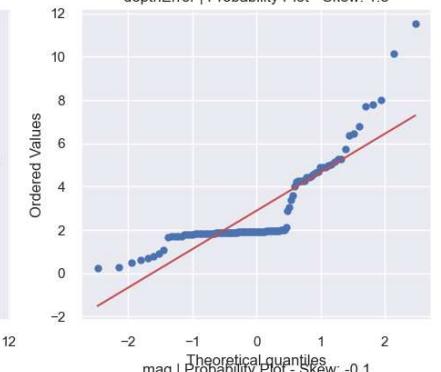
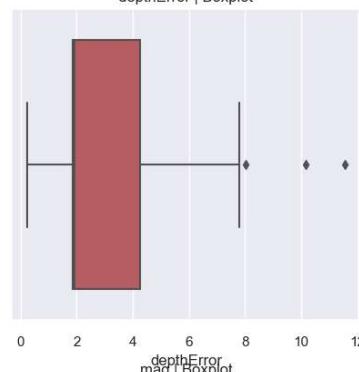
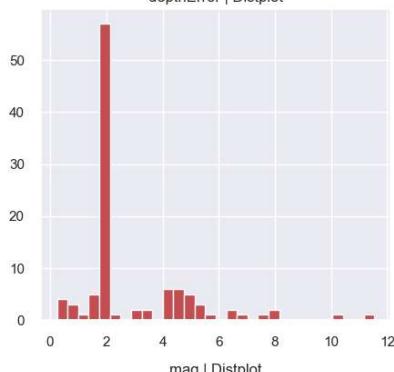
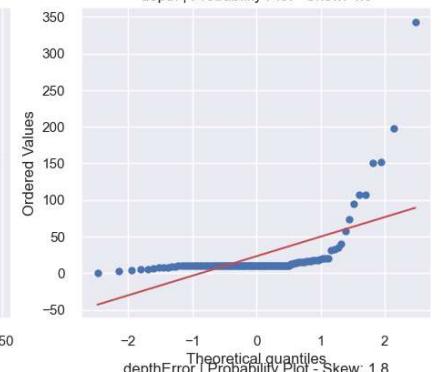
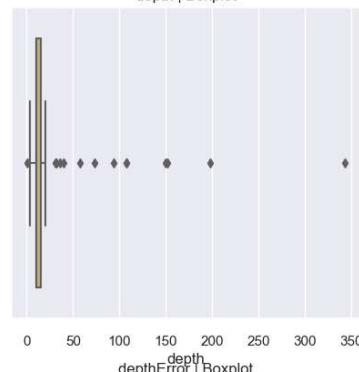
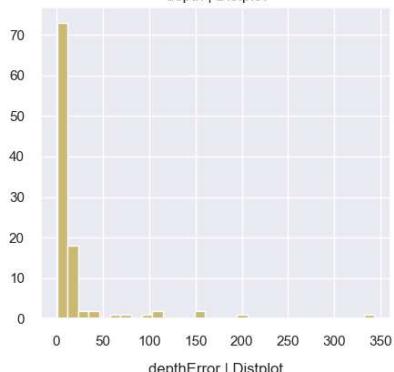
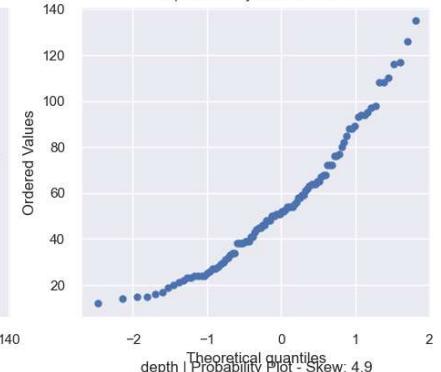
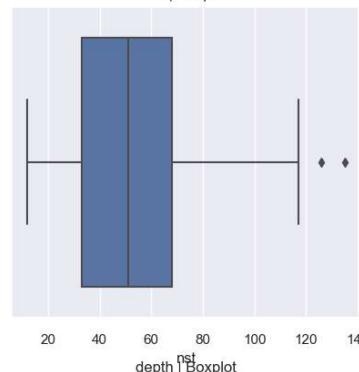
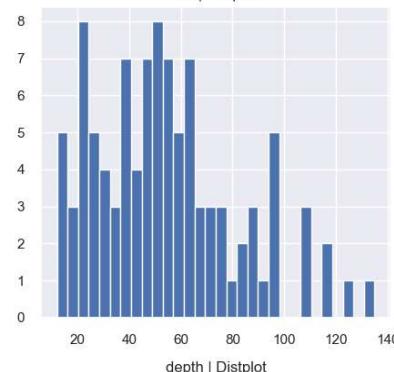
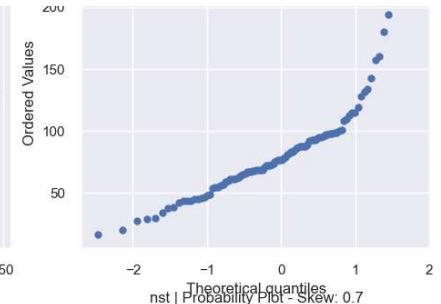
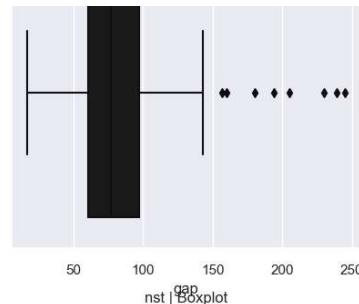
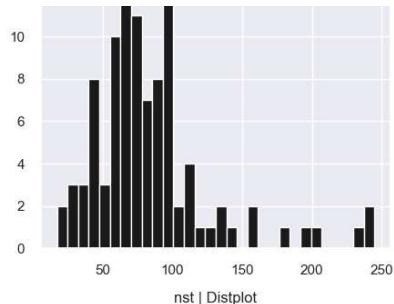




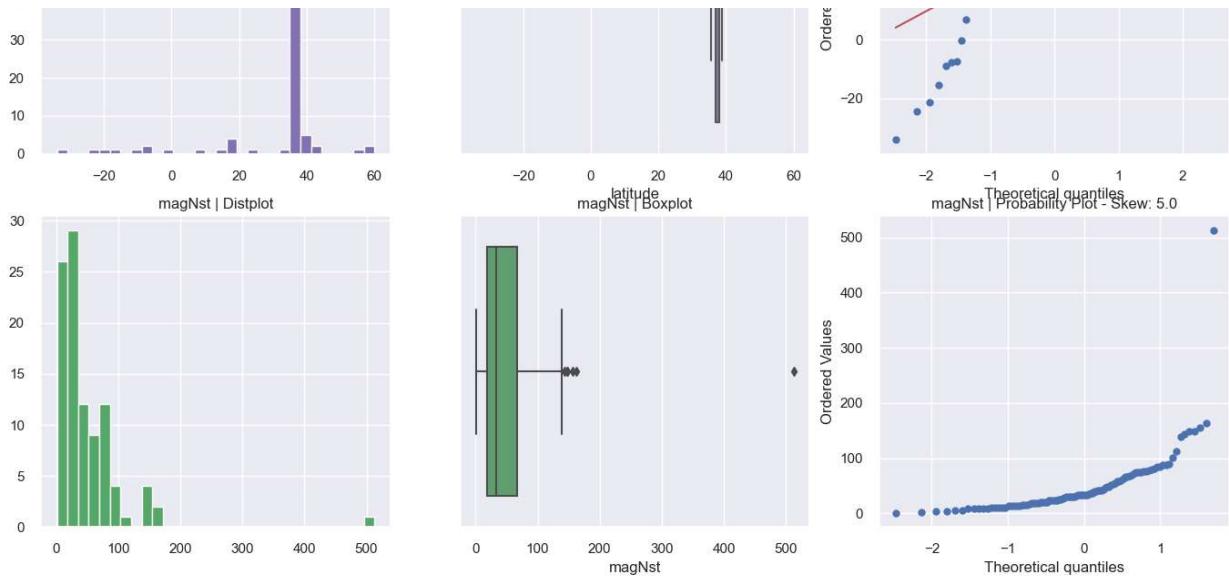
Turkey&Syria earthquake analysis



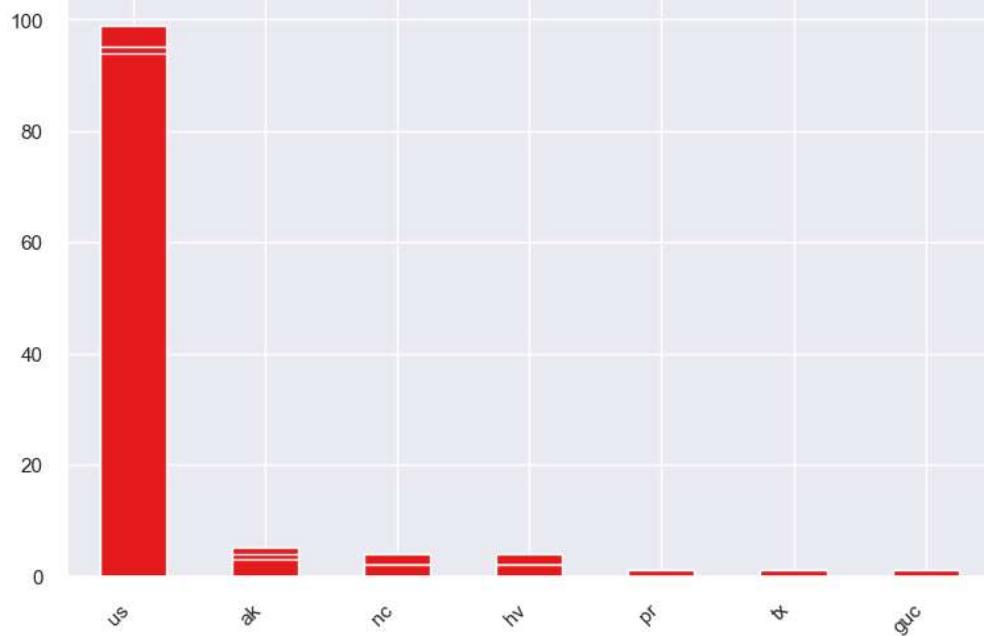
Turkey&Syria earthquake analysis



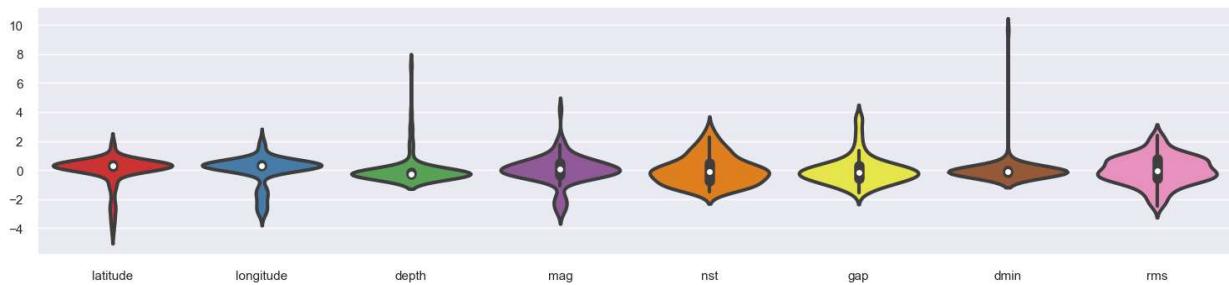
Turkey&Syria earthquake analysis



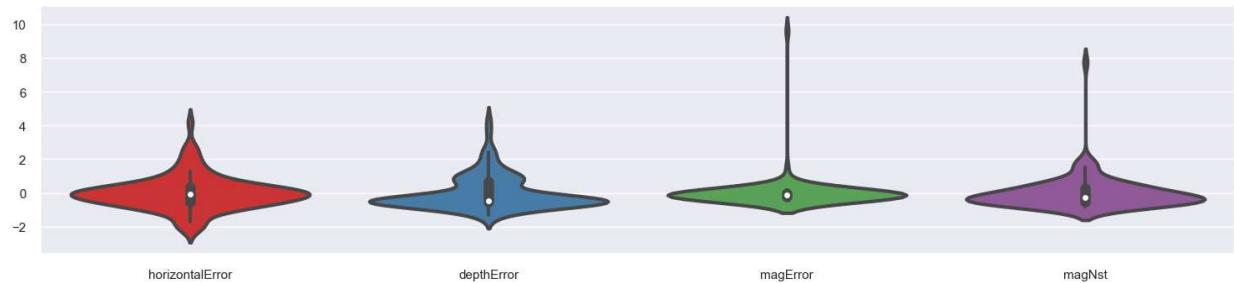
Histograms (KDE plots) of all Continuous Variables

Distribution of `magSource` (top 15 categories only)

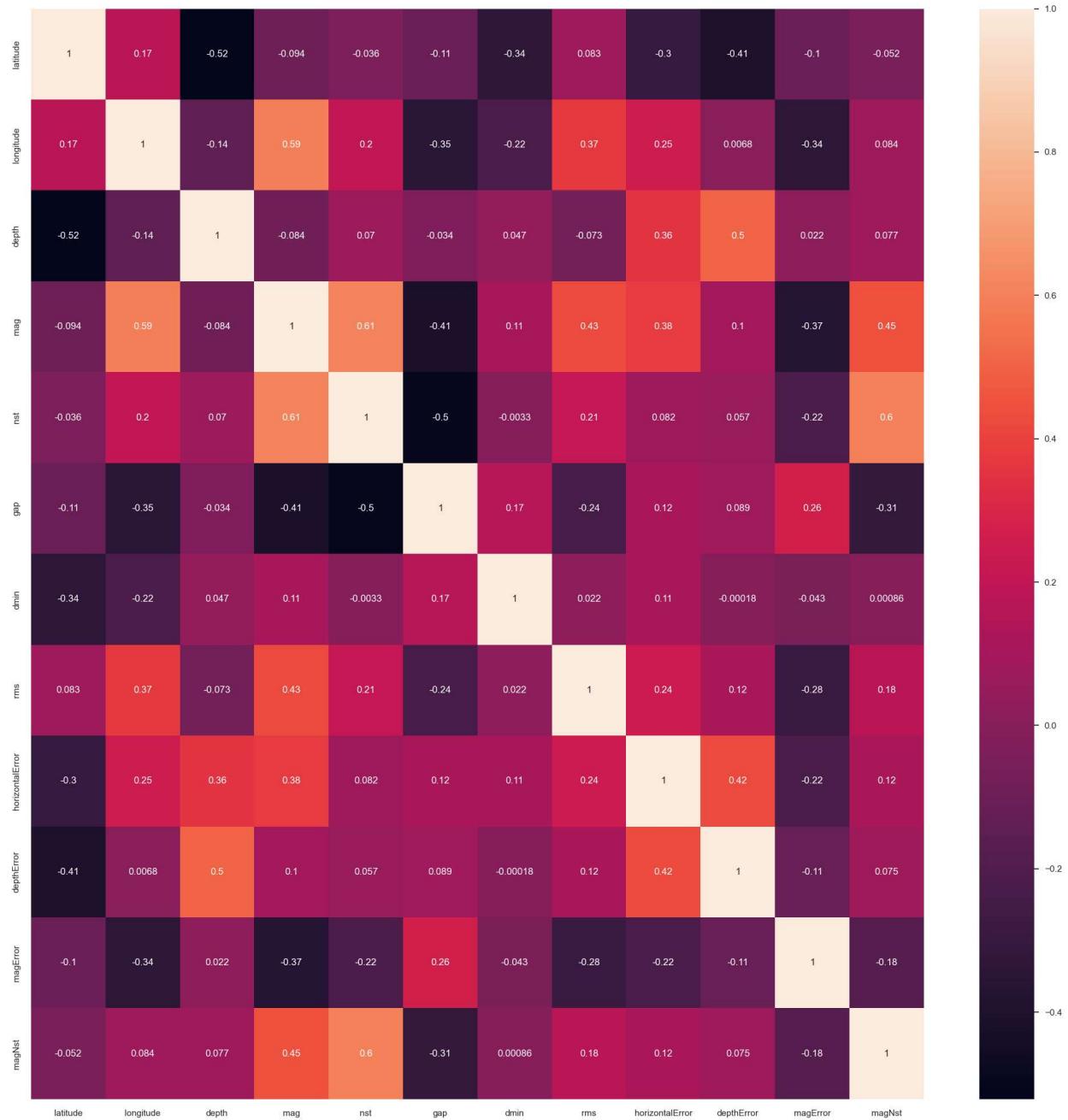
Violin Plot of all Continuous Variables



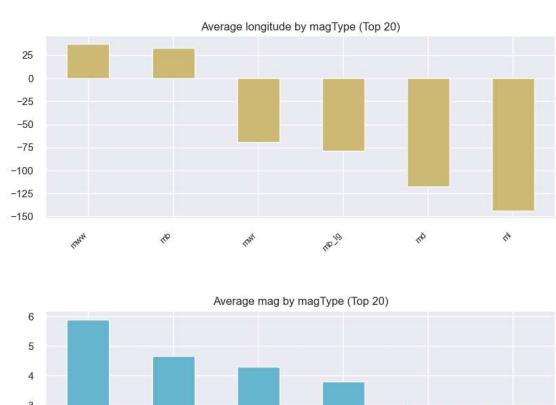
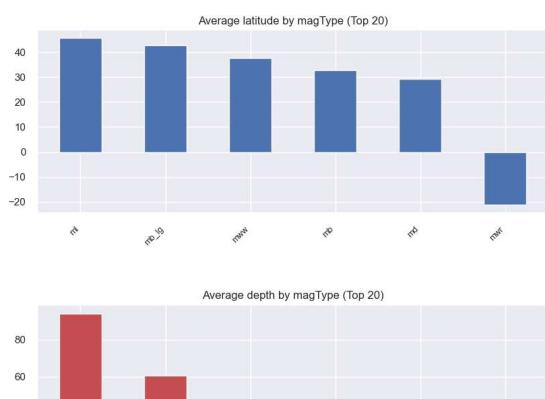
Violin Plot of all Continuous Variables



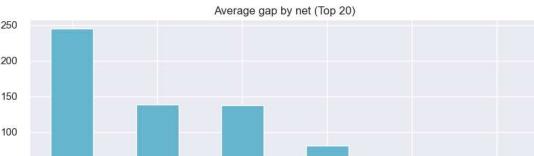
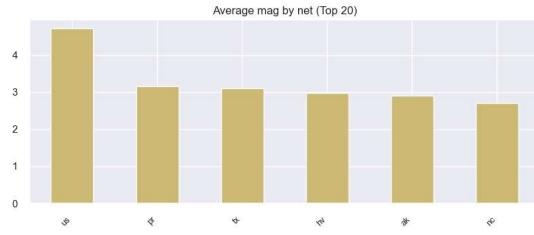
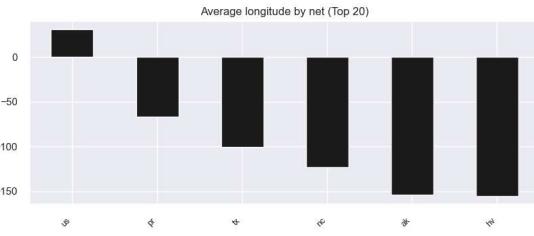
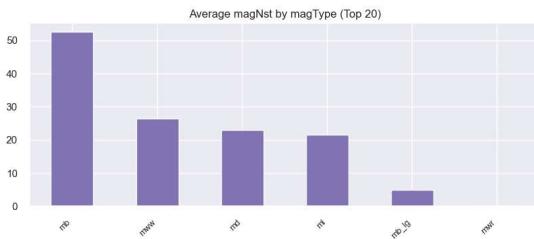
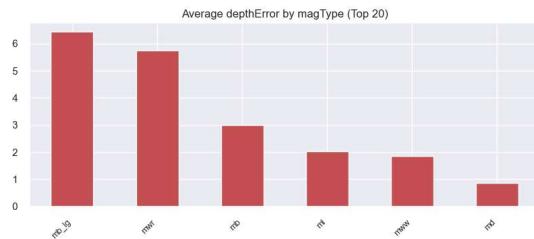
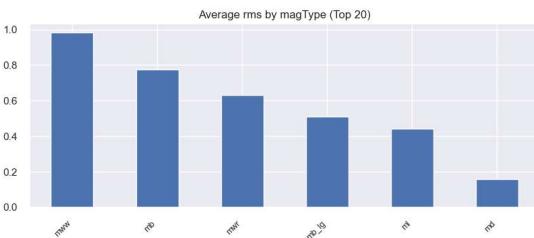
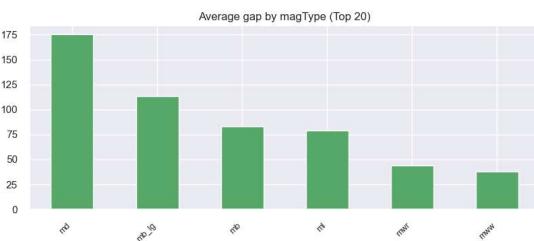
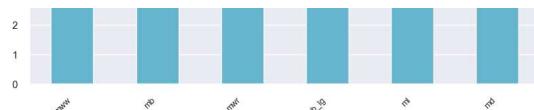
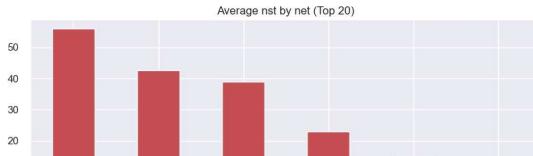
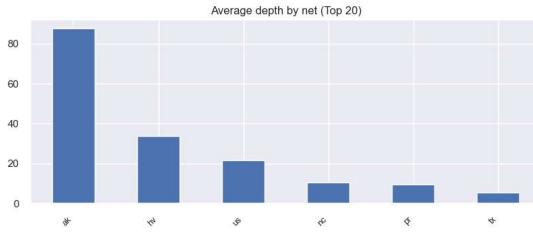
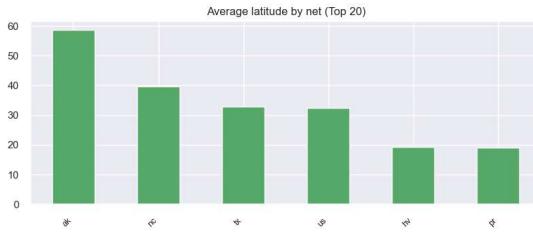
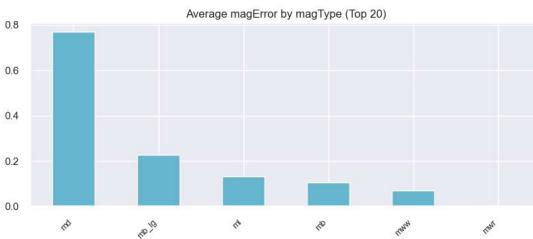
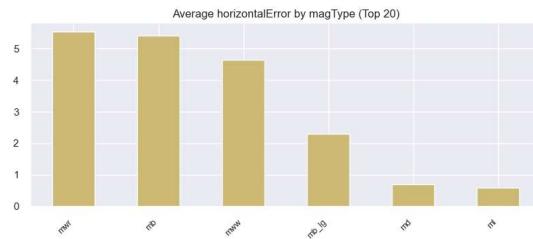
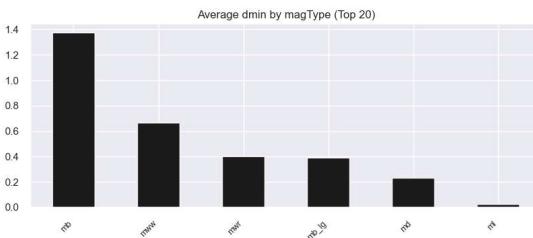
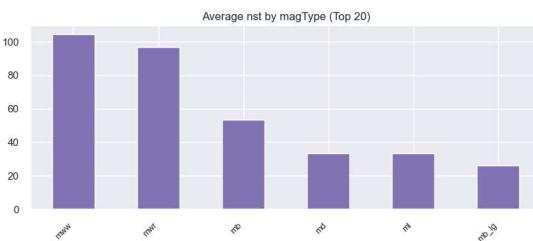
Heatmap of all Continuous Variables including target =



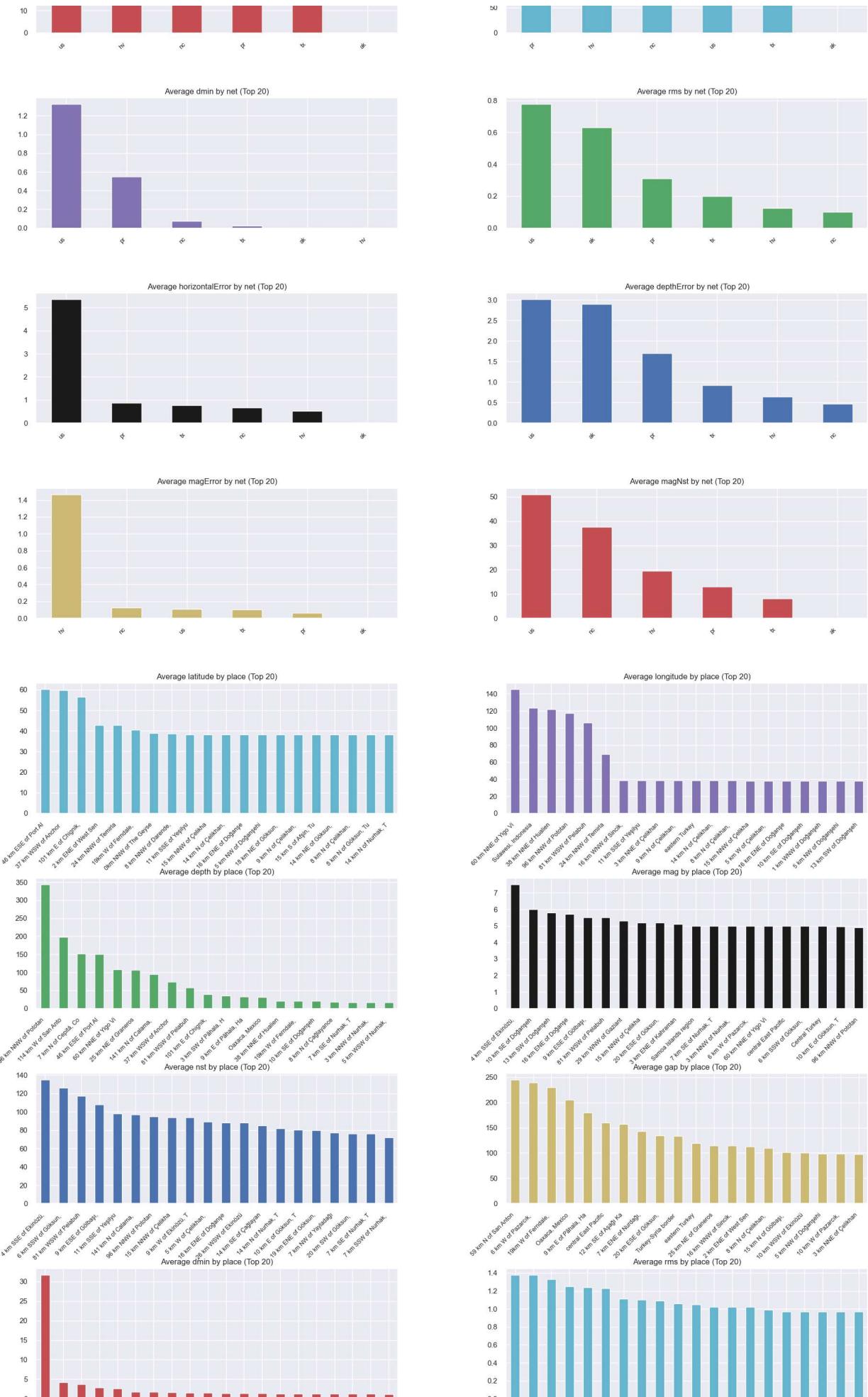
Bar plots for each Continuous by each Categorical variable



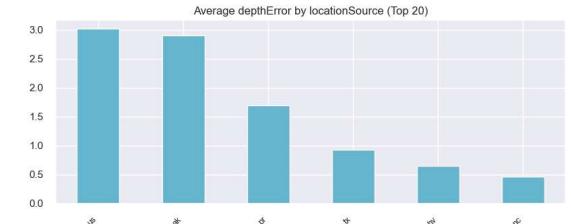
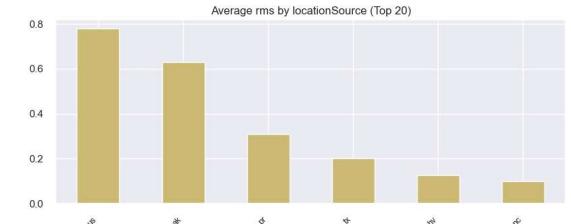
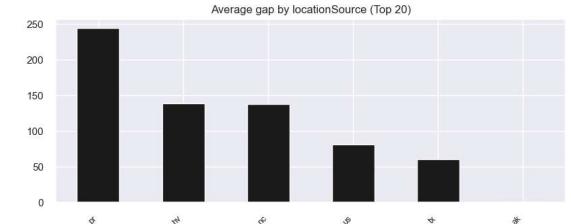
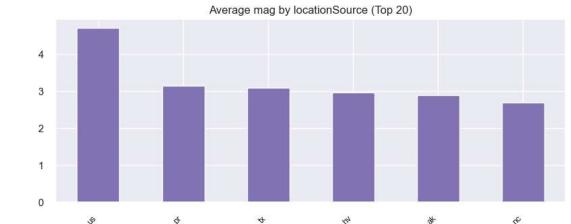
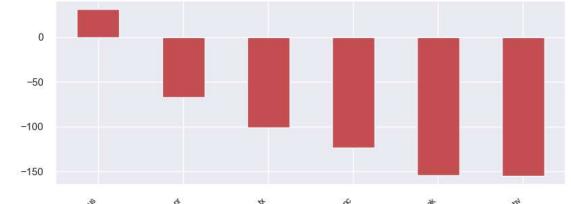
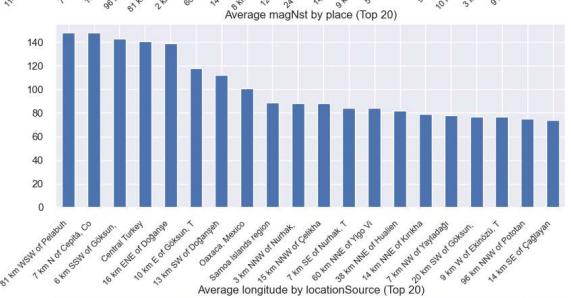
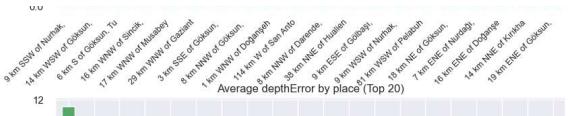
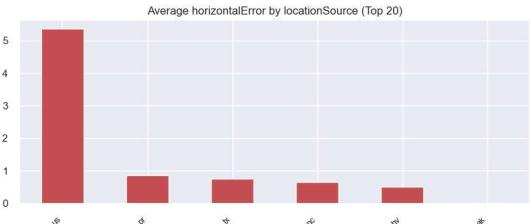
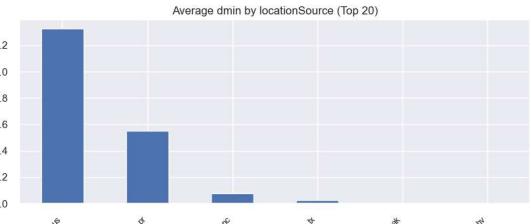
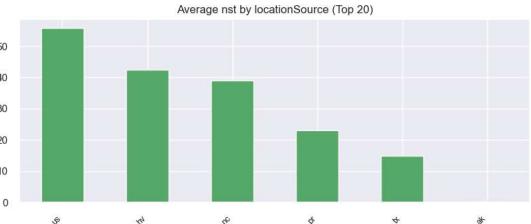
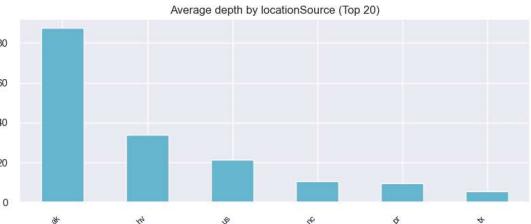
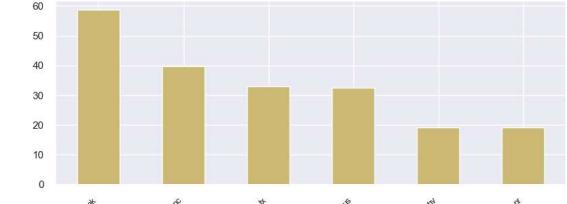
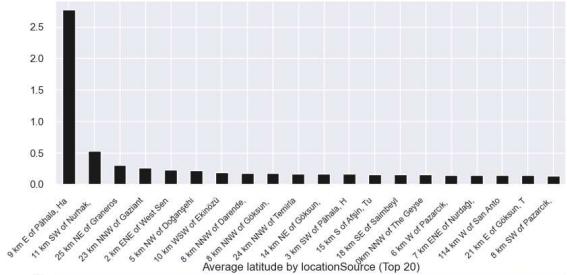
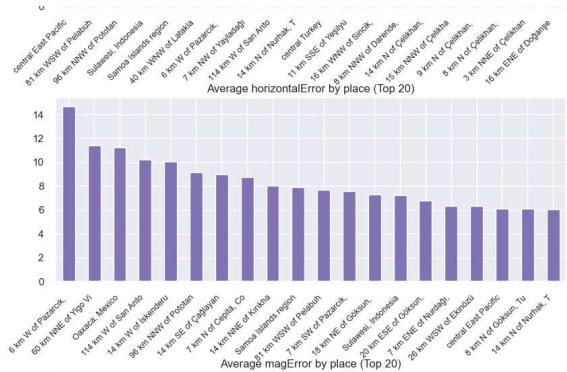
Turkey&Syria earthquake analysis

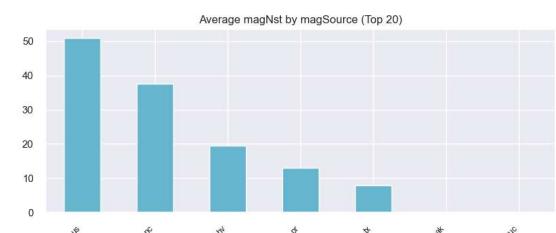
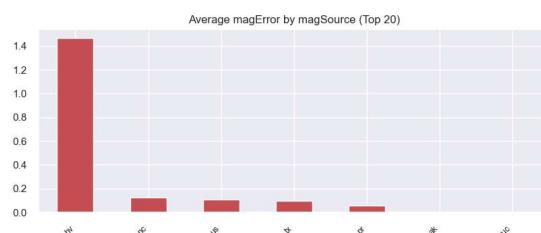
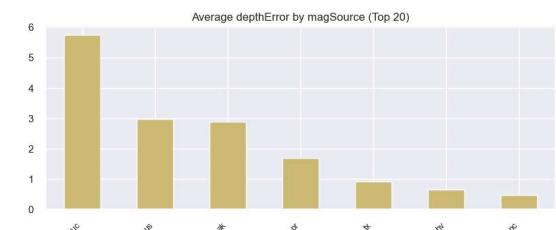
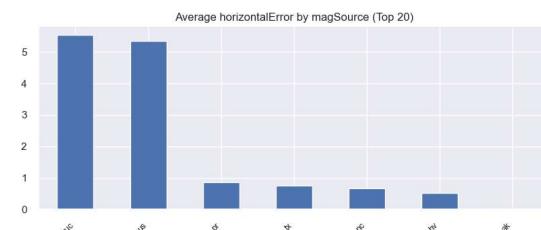
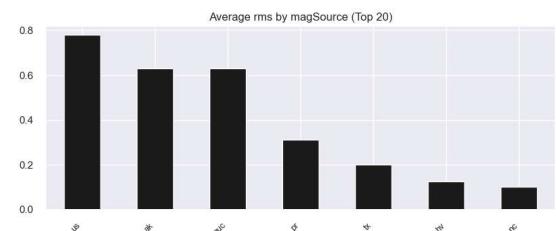
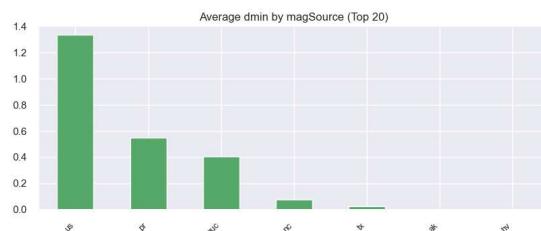
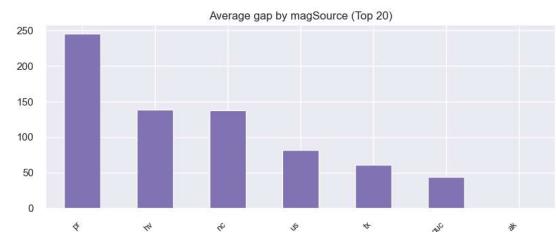
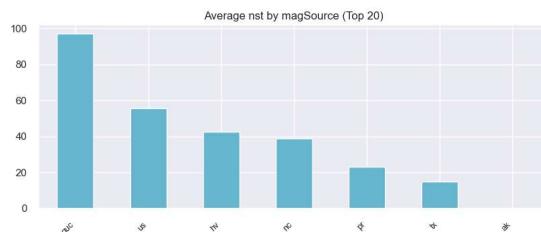
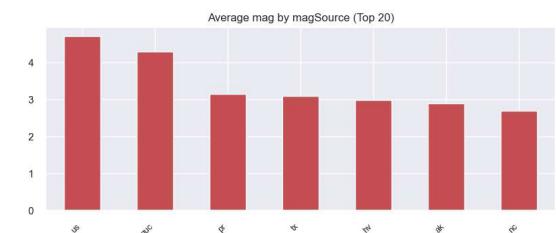
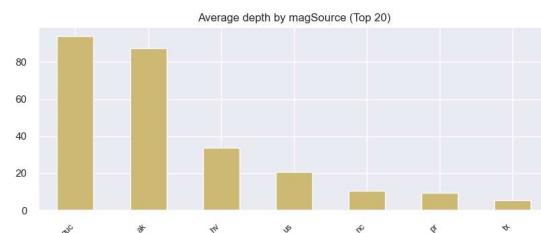
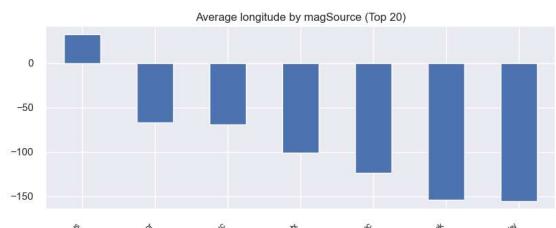
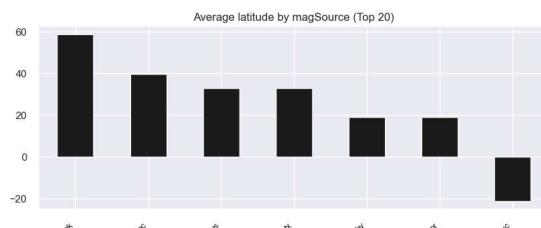
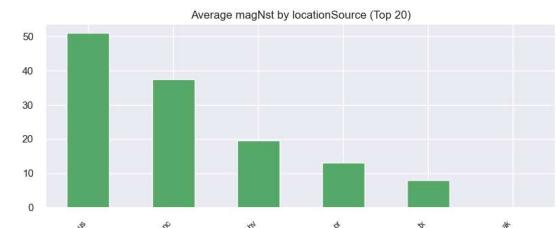
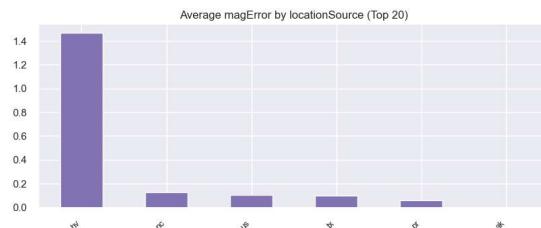


Turkey&Syria earthquake analysis

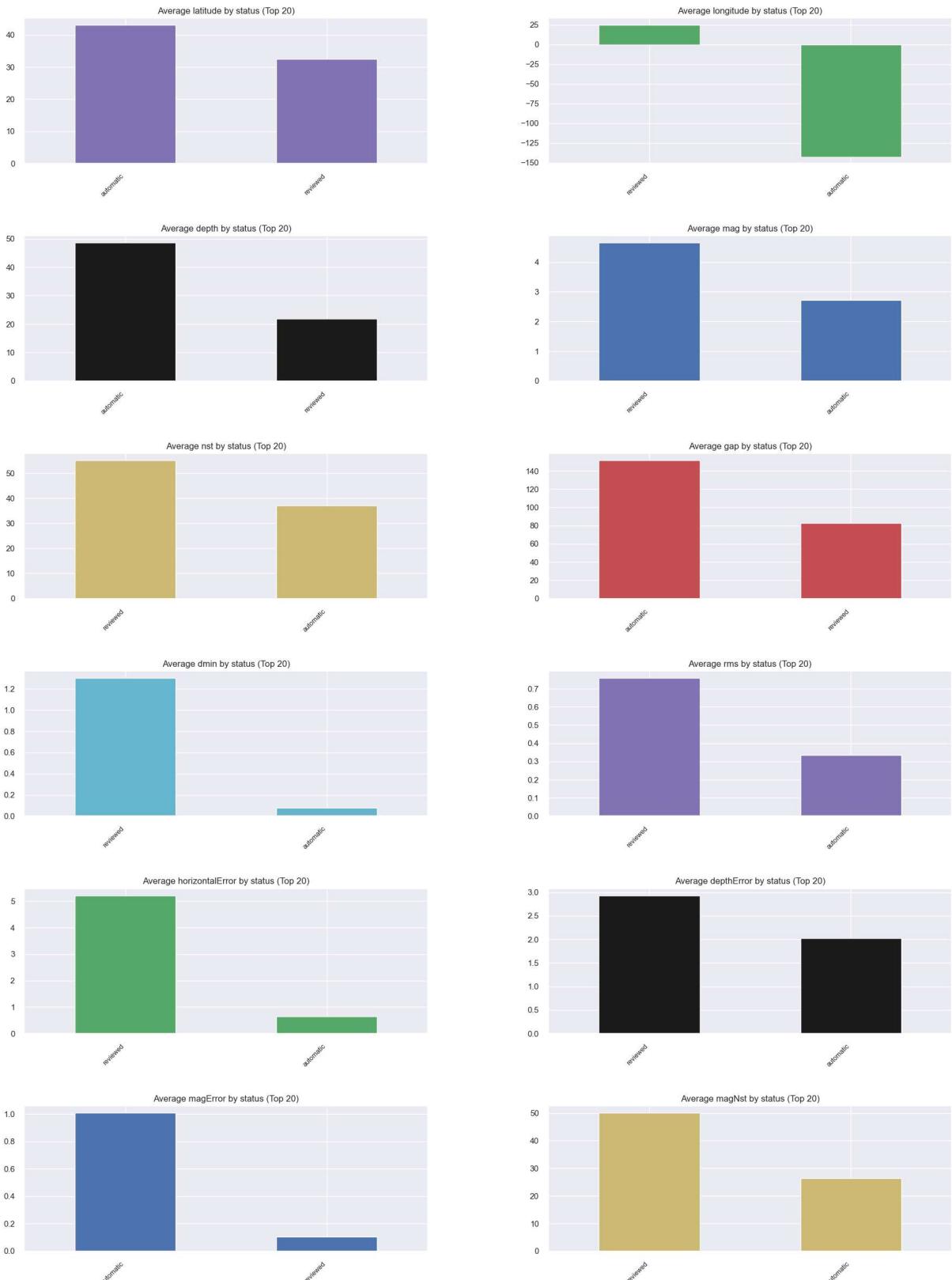


Turkey&Syria earthquake analysis





Turkey&Syria earthquake analysis



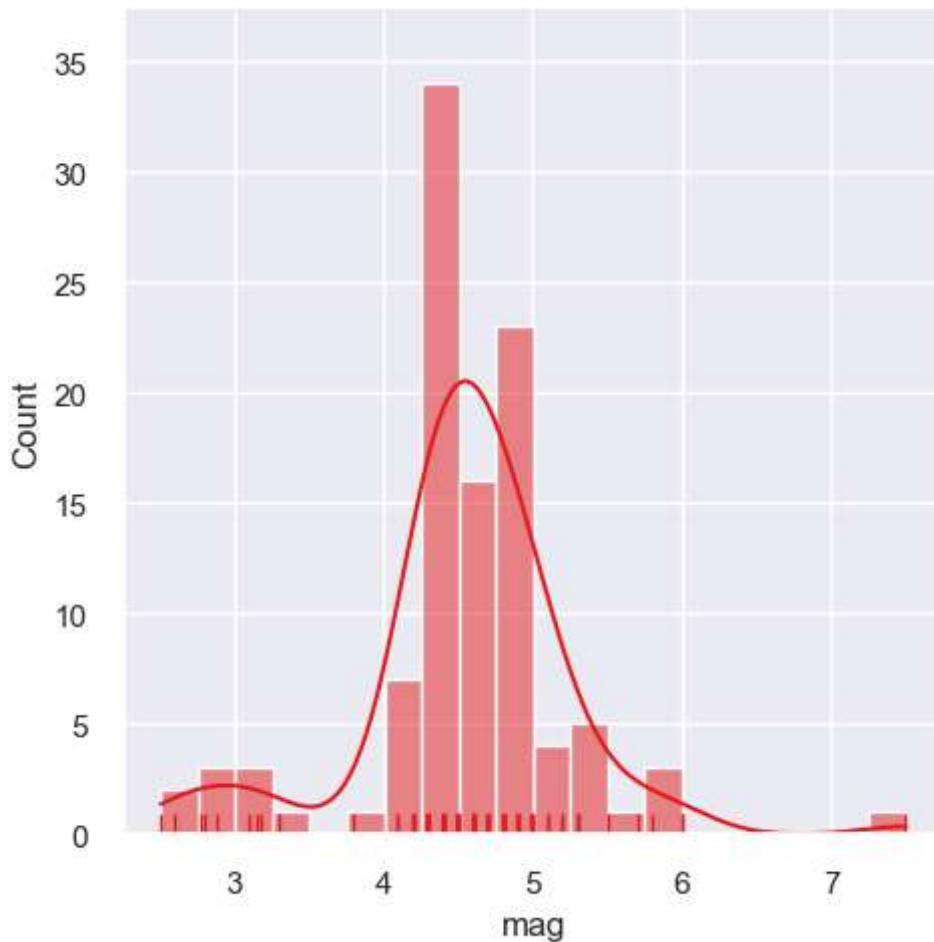
All Plots done

Time to run AutoViz = 21 seconds

#####
AUTO VISUALIZATION Completed #####
#####

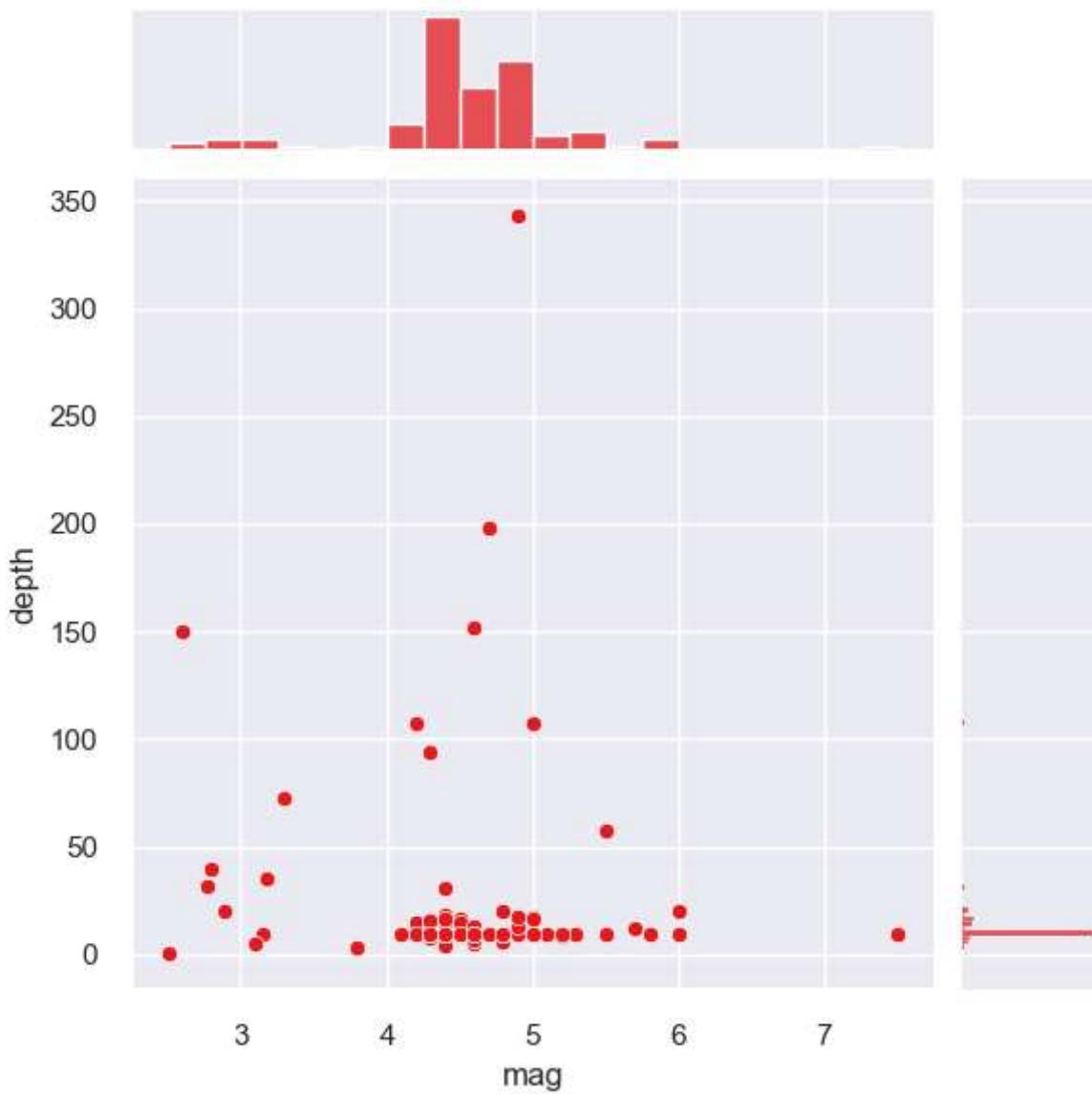
In [37]: `sns.displot(data['mag'], rug = True, kde = True)`

Out[37]: `<seaborn.axisgrid.FacetGrid at 0x1ebcd49cf10>`



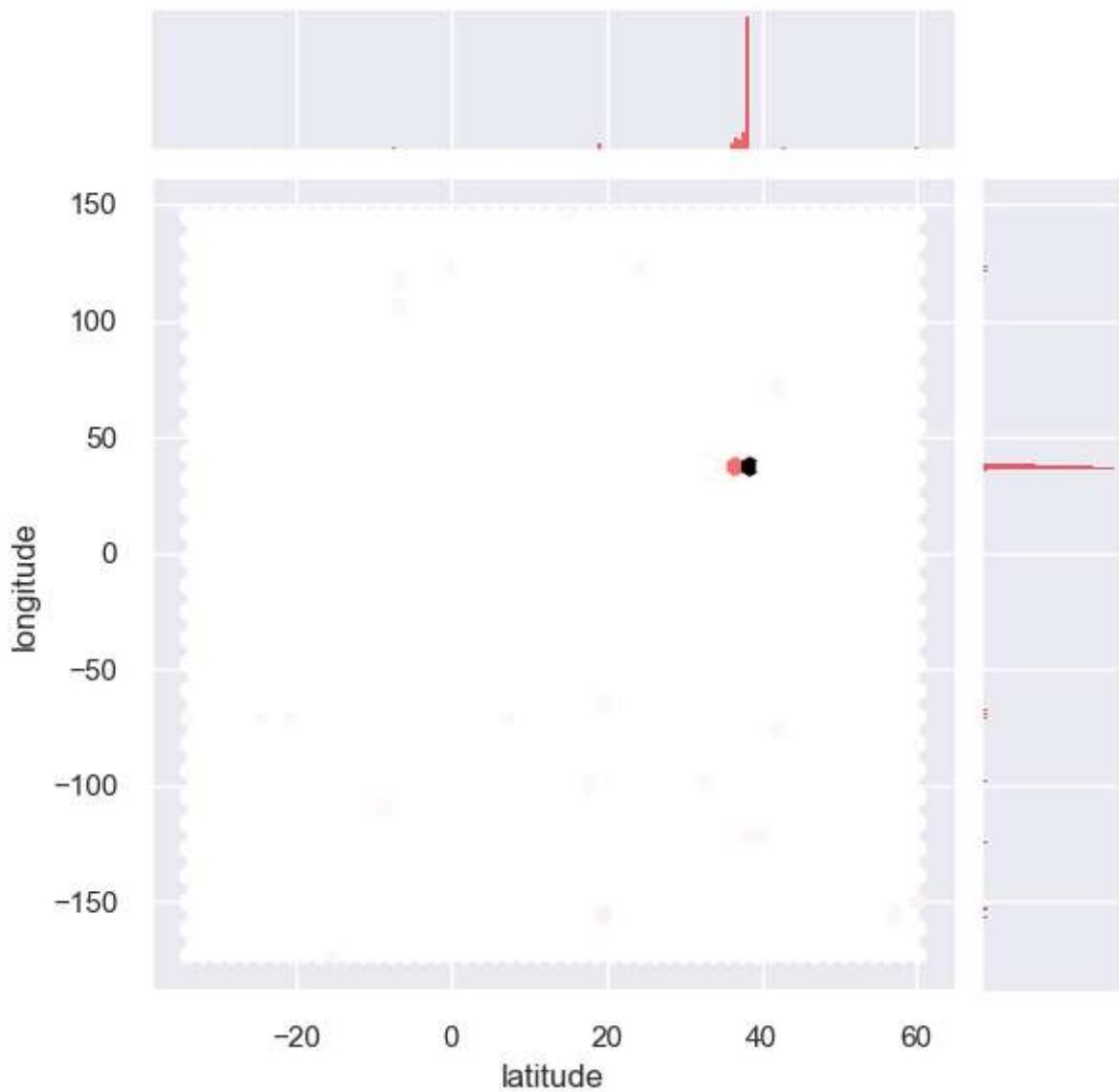
```
In [50]: sns.jointplot(data['mag'], data['depth'], data['type'])
```

```
Out[50]: <seaborn.axisgrid.JointGrid at 0x1ebd6cc7a30>
```



```
In [65]: sns.jointplot(data['latitude'], data['longitude'], kind = 'hex')
```

```
Out[65]: <seaborn.axisgrid.JointGrid at 0x1ebd76444c0>
```



In []: