In [26]:

```
import pandas as pd
df = pd.read_csv('merged_file.csv')
```

In [27]:

```
df.head(10)
```

Out[27]:

	DEPTH	DT	FACIES	FLD1	GR	NPHI	RHOB	ZLT
0	1295.9144	137.8066	NaN	NaN	61.3278	0.5643	2.1857	NaN
1	1296.0668	139.5873	0.0	NaN	61.9954	0.5611	2.1762	NaN
2	1296.2192	140.0185	0.0	NaN	63.5188	0.5630	2.1946	NaN
3	1296.3716	139.3474	0.0	NaN	64.9925	0.5677	2.1992	NaN
4	1296.5240	138.8638	0.0	NaN	65.6985	0.5743	2.1992	NaN
5	1296.6764	139.0847	0.0	NaN	65.1353	0.5844	2.2009	NaN
6	1296.8288	139.2288	0.0	NaN	63.4583	0.5984	2.2021	NaN
7	1296.9812	138.7143	0.0	NaN	61.7829	0.6146	2.2090	NaN
8	1297.1336	137.7427	0.0	NaN	60.9213	0.6306	2.2248	NaN
9	1297.2860	136.4887	0.0	NaN	60.5683	0.6386	2.2439	NaN

In [28]:

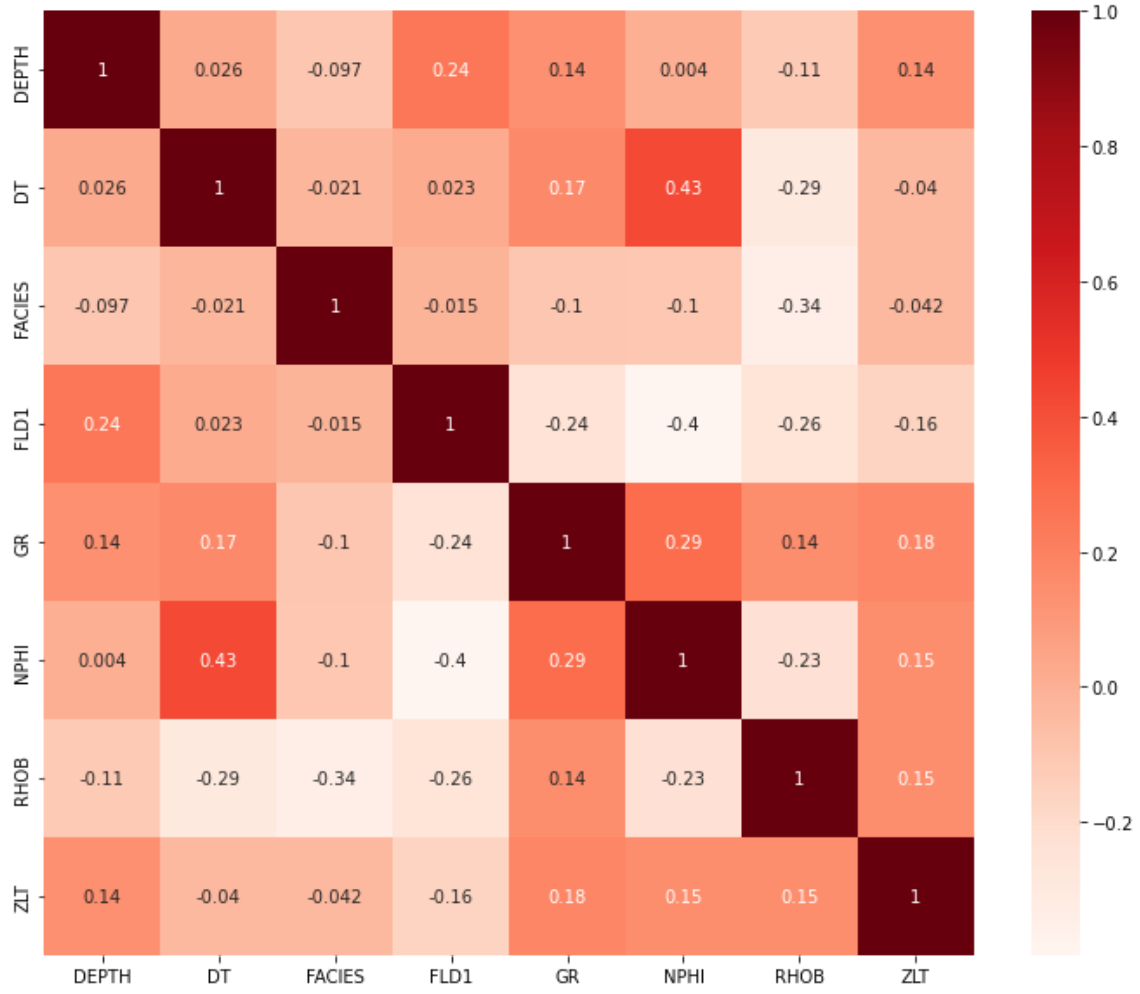
```
df = df.dropna().reset_index(drop=True)
```

In [29]:

```
import seaborn as sns
import matplotlib.pyplot as plt
```

In [30]:

```
plt.figure(figsize=(12,10))
cor = df.corr()
sns.heatmap(cor, annot=True, cmap=plt.cm.Reds)
plt.show()
```



In [31]:

```
#Correlation with output variable
cor_target = abs(cor["FACIES"])
#Selecting highly correlated features
relevant_features = cor_target[cor_target>0.1]
relevant_features
```

Out[31]:

```
FACIES    1.000000
GR        0.101856
NPHI      0.104378
RHOB      0.340015
Name: FACIES, dtype: float64
```

In [32]:

```
df = df.drop(['DEPTH' , 'DT' , 'ZLT' , 'FLD1'], axis=1)
df
```

Out[32]:

	FACIES	GR	NPHI	RHOB
0	3.0	31.5743	0.5045	1.7643
1	3.0	39.3396	0.4365	2.0439
2	0.0	46.5190	0.4037	2.2661
3	0.0	52.1829	0.3938	2.3546
4	0.0	56.4486	0.3974	2.3663
...	...	...	...	...
3958	0.0	74.6066	0.5261	2.4379
3959	0.0	76.7127	0.5439	2.4342
3960	0.0	77.0013	0.5283	2.4508
3961	0.0	72.7778	0.5135	2.4784
3962	0.0	68.5550	0.5175	2.4600

3963 rows × 4 columns

In [33]:

```
df.head(10)
```

Out[33]:

	FACIES	GR	NPHI	RHOB
0	3.0	31.5743	0.5045	1.7643
1	3.0	39.3396	0.4365	2.0439
2	0.0	46.5190	0.4037	2.2661
3	0.0	52.1829	0.3938	2.3546
4	0.0	56.4486	0.3974	2.3663
5	0.0	59.9263	0.4090	2.3813
6	0.0	63.7691	0.4256	2.4239
7	0.0	67.9563	0.4443	2.4968
8	0.0	69.9204	0.4626	2.5836
9	0.0	69.1420	0.4779	2.6326

In [35]:

```
df.to_csv('merged_file1.csv', index=False)
```