**Day 77- 90 days of Analytics : Missing Values in Pandas**

In today’s video, we looked some ways we can handle missing values with Pandas

The following were mentioned

-One way to deal with empty cells is to remove rows that contain empty cells. This is usually OK, since data sets can be very big, and removing a few rows will not have a big impact on the result. We use the **dropna()** method

import pandas as pd

df = pd.read\_csv('data.csv')

new\_df = df.dropna()

print(new\_df)

-We should note that by default, the dropna() method returns a new DataFrame, and will not change the original.

-If we want to change the original DataFrame, we use the inplace = True argument. df.dropna(inplace = True)

-The **thresh** property of the fillna() method permits us to delete rows with specific a number of missing values per column.

-Another way of dealing with empty cells is to insert a new value instead. This way we do not have to delete entire rows just because of some empty cells. The **fillna()** method allows us to replace empty cells with a value

import pandas as pd

df = pd.read\_csv('data\_store.csv')

new\_df = df.fillna(130)

print(new\_df)

-To only replace empty values for one column, specify the ***column name*** for the DataFrame. Example

import pandas as pd

df = pd.read\_csv('data\_store.csv')

new\_df = df["units\_sold"].fillna(0)

print(new\_df)

-We can replace empty values in various columns with different values by passing a dictionary to the fillna() method. Example

new\_df = df.fillna({

'visitors':0,

'units\_sold':0,

'sex':"Not provided"

})

print(new\_df)

-The replace() method permits us to replace specific values in a data frame. Example

import pandas as pd

import numpy as np

df = pd.read\_csv('store\_data2.csv')

df2 = df.replace(111111,value=np.nan)

print(df2)

Link to the YouTube Recording: <https://www.youtube.com/watch?v=L1w5QeoOlGc>

[#90daysofanalytics](https://www.linkedin.com/feed/hashtag/?keywords=90daysofanalytics&highlightedUpdateUrns=urn%3Ali%3Aactivity%3A7036754603295539200) [#community](https://www.linkedin.com/feed/hashtag/?keywords=community&highlightedUpdateUrns=urn%3Ali%3Aactivity%3A7036754603295539200) [#dataanalysis](https://www.linkedin.com/feed/hashtag/?keywords=dataanalysis&highlightedUpdateUrns=urn%3Ali%3Aactivity%3A7036754603295539200) [#dataanalyst](https://www.linkedin.com/feed/hashtag/?keywords=dataanalyst&highlightedUpdateUrns=urn%3Ali%3Aactivity%3A7036754603295539200) #microsoft #msexcel #SQL #powerbi #pythonprogramming #numpy