fb

April 29, 2025

[1]:

*# Import necessary libraries*

**import pandas as pd import numpy as np**

# Loading the Facebook Metrics Dataset

[5]:

*# Load the dataset*

fb\_data = pd.read\_csv('dataset\_facebook.csv') print(fb\_data)

*# Display the first few rows*

print("Facebook Metrics Dataset Shape:", fb\_data.shape) fb\_data.head()

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | pagetotallikes | type | category | postmonth | postweekday | posthour | paid | \ |
| 0 | 139441 | 2 | 2 | 12 | 4 | 3 | 0 |  |
| 1 | 139441 | 3 | 2 | 12 | 3 | 10 | 0 |  |
| 2 | 139441 | 2 | 3 | 12 | 3 | 3 | 0 |  |
| 3 | 139441 | 2 | 2 | 12 | 2 | 10 | 1 |  |
| 4  ..  367 | 139441  …  85093 | 2  …  2 | 2  …  3 | 12  …  1 | 2  … …  7 | 3  …  10 | 0  0 |  |
| 368 | 85093 | 2 | 3 | 1 | 7 | 2 | 0 |  |
| 369 | 81370 | 2 | 2 | 1 | 5 | 8 | 0 |  |
| 370 | 81370 | 2 | 1 | 1 | 5 | 2 | 0 |  |
| 371 | 81370 | 2 | 3 | 1 | 4 | 11 | 0 |  |
|  | lifetimeposttotalreach | | lifetimeposttotalimpressions | | | \ | | |
| 0 | 2752 | | 5091 | | |  | | |
| 1 | 10460 | | 19057 | | |  | | |
| 2 | 2413 | | 4373 | | |  | | |
| 3 | 50128 | | 87991 | | |  | | |
| 4  ..  367 | 7244  …  5400 | | 13594  …  9218 | | |  | | |
| 368 | 4684 | | 7536 | | |  | | |
| 369 | 3480 | | 6229 | | |  | | |
| 370 | 3778 | | 7216 | | |  | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 371 | 4156 | 7564 | | | | |
|  | lifetimeengagedusers | lifetimepostconsumers | | lifetimepostconsumptions | | \ |
| 0 | 178 | 109 | | 159 | |  |
| 1 | 1457 | 1361 | | 1674 | |  |
| 2 | 177 | 113 | | 154 | |  |
| 3 | 2211 | 790 | | 1119 | |  |
| 4  ..  367 | 671  …  810 | 410  …  756 | | 580  …  1003 | |  |
| 368 | 733 | 708 | | 985 | |  |
| 369 | 537 | 508 | | 687 | |  |
| 370 | 625 | 572 | | 795 | |  |
| 371 | 626 | 574 | | 832 | |  |
| 0 | lifetimepostimpressionsbypeoplewhohavelikedyourpage  3078 | | | | \ | |
| 1 | 11710 | | | |  | |
| 2 | 2812 | | | |  | |
| 3 | 61027 | | | |  | |
| 4  ..  367 | 6228  …  5654 | | | |  | |
| 368 | 4750 | | | |  | |
| 369 | 3961 | | | |  | |
| 370 | 4742 | | | |  | |
| 371 | 4534 | | | |  | |
| 0 | lifetimepostreachbypeoplewholikeyourpage  1640 | | \ | | | |
| 1 | 6112 | |  | | | |
| 2 | 1503 | |  | | | |
| 3 | 32048 | |  | | | |
| 4  ..  367 | 3200  …  3230 | |  | | | |
| 368 | 2876 | |  | | | |
| 369 | 2104 | |  | | | |
| 370 | 2388 | |  | | | |
| 371 | 2452 | |  | | | |

lifetimepeoplewhohavelikedyourpageandengagedwithyourpost comment like \

0 119 4 79

1 1108 5 130

2 132 0 66

3 1386 58 1572

4 396 19 325

.. … … …

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 367 | |  |  | | | 422 | | 10 | | 125 |
| 368 | |  |  | | | 392 | | 5 | | 53 |
| 369 | |  |  | | | 301 | | 0 | | 53 |
| 370 | |  |  | | | 363 | | 4 | | 93 |
| 371 | |  |  | | | 370 | | 7 | | 91 |
|  | | share | totalinteractions | | |  | |  | |  |
| 0 | | 17 | 100 | | |  | |  | |  |
| 1 | | 29 | 164 | | |  | |  | |  |
| 2 | | 14 | 80 | | |  | |  | |  |
| 3 | | 147 | 1777 | | |  | |  | |  |
| 4 | | 49 | 393 | | |  | |  | |  |
| .. | | … | … | | |  | |  | |  |
| 367 | | 41 | 176 | | |  | |  | |  |
| 368 | | 26 | 84 | | |  | |  | |  |
| 369 | | 22 | 75 | | |  | |  | |  |
| 370 | | 18 | 115 | | |  | |  | |  |
| 371 | | 38 | 136 | | |  | |  | |  |
| [372 | | rows x | 19 columns] | | |  | |  | |  |
| Facebook Metrics Dataset Shape: (372, 19) | | | | | | | | | | |
| [5]: |  | pagetotallikes | | type | category | postmonth | postweekday | posthour | paid \ | |
|  | 0 | 139441 | | 2 | 2 | 12 | 4 | 3 | 0 | |
|  | 1 | 139441 | | 3 | 2 | 12 | 3 | 10 | 0 | |
|  | 2 | 139441 | | 2 | 3 | 12 | 3 | 3 | 0 | |
|  | 3 | 139441 | | 2 | 2 | 12 | 2 | 10 | 1 | |
|  | 4 | 139441 | | 2 | 2 | 12 | 2 | 3 | 0 | |

|  |  |  |  |
| --- | --- | --- | --- |
| lifetimeposttotalreach | lifetimeposttotalimpressions | lifetimeengagedusers | \ |
| 0 2752 | 5091 | 178 |  |
| 1 10460 | 19057 | 1457 |  |
| 2 2413 | 4373 | 177 |  |
| 3 50128 | 87991 | 2211 |  |
| 4 7244 | 13594 | 671 |  |

lifetimepostconsumers lifetimepostconsumptions \

0 109 159

1 1361 1674

2 113 154

3 790 1119

4 410 580

lifetimepostimpressionsbypeoplewhohavelikedyourpage \

0 3078

1 11710

2 2812

|  |  |  |
| --- | --- | --- |
| 3 | | 61027 |
| 4 | | 6228 |
| lifetimepostreachbypeoplewholikeyourpage \ | |  |
| 0 | 1640 | |
| 1 | 6112 | |
| 2 | 1503 | |
| 3 | 32048 | |
| 4 | 3200 | |

lifetimepeoplewhohavelikedyourpageandengagedwithyourpost comment like \

0 119 4 79

1 1108 5 130

2 132 0 66

3 1386 58 1572

4 396 19 325

share totalinteractions

|  |  |  |
| --- | --- | --- |
| 0 | 17 | 100 |
| 1 | 29 | 164 |
| 2 | 14 | 80 |
| 3 | 147 | 1777 |
| 4 | 49 | 393 |

[7]:

*# 1. Create a subset based on column selection* subset\_columns = fb\_data[['comment', 'like', 'share']] print("Subset by columns shape:", subset\_columns.shape) subset\_columns.head()

# Create Data Subsets

Subset by columns shape: (372, 3)

|  |  |  |  |
| --- | --- | --- | --- |
| [7]: | comment | like | share |
| 0 | 4 | 79 | 17 |
| 1 | 5 | 130 | 29 |
| 2 | 0 | 66 | 14 |
| 3 | 58 | 1572 | 147 |
| 4 | 19 | 325 | 49 |

[9]:

*# 2. Create a subset based on row filtering*

subset\_high\_likes = fb\_data[fb\_data['like'] > 200] print("Subset by high likes count:", subset\_high\_likes.shape) subset\_high\_likes.head()

Subset by high likes count: (75, 19)

[9]: pagetotallikes type category postmonth postweekday posthour paid \

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 139441 | 2 | 2 | 12 | 2 | 10 | 1 |
| 4 | 139441 | 2 | 2 | 12 | 2 | 3 | 0 |
| 6 | 139441 | 2 | 3 | 12 | 1 | 3 | 1 |
| 7 | 139441 | 2 | 3 | 12 | 7 | 9 | 1 |
| 10 | 139441 | 3 | 2 | 12 | 5 | 10 | 0 |
| lifetimeposttotalreach lifetimeposttotalimpressions \ | | | | | | | |

|  |  |  |
| --- | --- | --- |
| 3 | 147 | 1777 |
| 4 | 49 | 393 |
| 6 | 27 | 279 |
| 7 | 14 | 339 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 3 | 50128 | 87991 | | | | |
| 4 | 7244 | 13594 | | | | |
| 6 | 11692 | 19479 | | | | |
| 7 | 13720 | 24137 | | | | |
| 10 | 21744 | 42334 | | | | |
|  | lifetimeengagedusers | lifetimepostconsumers | lifetimepostconsumptions | | | \ |
| 3 | 2211 | 790 | 1119 | | |  |
| 4 | 671 | 410 | 580 | | |  |
| 6 | 481 | 265 | 364 | | |  |
| 7 | 537 | 232 | 305 | | |  |
| 10 | 4258 | 4100 | 4540 | | |  |
| lifetimepostimpressionsbypeoplewhohavelikedyourpage \ | | | | | | |
| 3 | 61027 | | | | | |
| 4 | 6228 | | | | | |
| 6 | 15432 | | | | | |
| 7 | 19728 | | | | | |
| 10 | 37849 | | | | | |
| lifetimepostreachbypeoplewholikeyourpage \ | | | | | | |
| 3 | 32048 | | | | | |
| 4 | 3200 | | | | | |
| 6 | 9328 | | | | | |
| 7 | 11056 | | | | | |
| 10 | 18952 | | | | | |
|  | lifetimepeoplewhohavelikedyourpageandengagedwithyourpost | | | comment | like | \ |
| 3 | 1386 | | | 58 | 1572 |  |
| 4 | 396 | | | 19 | 325 |  |
| 6 | 379 | | | 3 | 249 |  |
| 7 | 422 | | | 0 | 325 |  |
| 10 | 3798 | | | 0 | 233 |  |
| share totalinteractions | | | | | | |

10 19 252

[11]:

*# 3. Create a subset using iloc (integer-position based indexing)*

subset\_iloc = fb\_data.iloc[10:20, 2:6]

print("Subset using iloc shape:", subset\_iloc.shape) subset\_iloc

Subset using iloc shape: (10, 4)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| [11]: | category | postmonth | postweekday | posthour |
| 10 | 2 | 12 | 5 | 10 |
| 11 | 2 | 12 | 5 | 10 |
| 12 | 2 | 12 | 5 | 10 |
| 13 | 2 | 12 | 5 | 3 |
| 14 | 2 | 12 | 4 | 5 |
| 15 | 2 | 12 | 3 | 10 |
| 16 | 3 | 12 | 3 | 3 |
| 17 | 1 | 12 | 2 | 12 |
| 18 | 3 | 12 | 2 | 3 |
| 19 | 3 | 12 | 1 | 11 |

[13]:

*# 4. Create a subset using loc (label-based indexing)*

fb\_data\_reset = fb\_data.reset\_index()

subset\_loc = fb\_data\_reset.loc[5:15, ['comment', 'like']] print("Subset using loc shape:", subset\_loc.shape) subset\_loc

Subset using loc shape: (11, 2)

[13]: comment like

|  |  |  |
| --- | --- | --- |
| 5 | 1 | 152 |
| 6 | 3 | 249 |
| 7 | 0 | 325 |
| 8 | 0 | 161 |
| 9 | 3 | 113 |
| 10 | 0 | 233 |
| 11 | 0 | 88 |
| 12 | 0 | 90 |
| 13 | 5 | 137 |
| 14 | 2 | 577 |
| 15 | 4 | 86 |

# Merge Data

[15]:

*# Create two dataframes to demonstrate merging # First dataframe*

df1 = fb\_data[["like"]].copy() df1.reset\_index(inplace=**True**)

*# Second dataframe*

df2 = fb\_data[["share"]].copy() df2.reset\_index(inplace=**True**)

print("DataFrame 1 shape:", df1.shape) print("DataFrame 2 shape:", df2.shape)

*# Show sample of both dataframes* print("**\n**DataFrame 1 sample:") display(df1.head(3)) print("**\n**DataFrame 2 sample:") display(df2.head(3))

DataFrame 1 shape: (372, 2)

DataFrame 2 shape: (372, 2)

DataFrame 1 sample: index like

0 0 79

1 1 130

2 2 66

[17]:

DataFrame 2 sample: index share

0 0 17

1 1 29

2 2 14

*# 1. Inner join (only keep rows found in both dataframes)*

merged\_inner = pd.merge(df1, df2, how='inner', suffixes=('\_profile',␣

𝗌'\_performance'))

print("Inner join shape:", merged\_inner.shape) merged\_inner.head()

Inner join shape: (372, 3)

|  |  |  |  |
| --- | --- | --- | --- |
| [17]: | index | like | share |
| 0 | 0 | 79 | 17 |
| 1 | 1 | 130 | 29 |
| 2 | 2 | 66 | 14 |
| 3 | 3 | 1572 | 147 |
| 4 | 4 | 325 | 49 |

[19]:

*# 2. Left join (keep all rows from df1)*

merged\_left = pd.merge(df1, df2, how='left', suffixes=('\_profile',␣

𝗌'\_performance'))

print("Left join shape:", merged\_left.shape) merged\_left.head()

Left join shape: (372, 3)

|  |  |  |  |
| --- | --- | --- | --- |
| [19]: | index | like | share |
| 0 | 0 | 79 | 17 |
| 1 | 1 | 130 | 29 |
| 2 | 2 | 66 | 14 |
| 3 | 3 | 1572 | 147 |
| 4 | 4 | 325 | 49 |

[21]:

*# 3. Right join (keep all rows from df2)*

merged\_right = pd.merge(df1, df2, how='right', suffixes=('\_profile',␣

𝗌'\_performance'))

print("Right join shape:", merged\_right.shape) merged\_right.head()

Right join shape: (372, 3)

|  |  |  |  |
| --- | --- | --- | --- |
| [21]: | index | like | share |
| 0 | 0 | 79 | 17 |
| 1 | 1 | 130 | 29 |
| 2 | 2 | 66 | 14 |
| 3 | 3 | 1572 | 147 |
| 4 | 4 | 325 | 49 |

[23]:

*# 4. Outer join (keep all rows from both dataframes)*

merged\_outer = pd.merge(df1, df2, how='outer', suffixes=('\_profile',␣

𝗌'\_performance'))

print("Outer join shape:", merged\_outer.shape) merged\_outer.head()

Outer join shape: (372, 3)

|  |  |  |  |
| --- | --- | --- | --- |
| [23]: | index | like | share |
| 0 | 0 | 79 | 17 |
| 1 | 1 | 130 | 29 |
| 2 | 2 | 66 | 14 |
| 3 | 3 | 1572 | 147 |
| 4 | 4 | 325 | 49 |

[26]:

*# 1. Sort by a single column (ascending)*

sorted\_likes = fb\_data.sort\_values(by='like')

# Sort Data

print("Sorted by Page total likes (ascending):") sorted\_likes[['like']].head(30)

Sorted by Page total likes (ascending):

[26]: like

76 0

290 0

314 0

21 0

100 0

120 1

301 2

295 2

302 3

121 3

126 3

305 4

114 4

304 4

125 4

354 5

303 6

299 6

118 6

117 7

128 7

116 7

129 7

122 7

127 8

124 9

130 11

49 12

358 12

103 13

[28]:

*# 2. Sort by a single column (descending)*

sorted\_likes\_desc = fb\_data.sort\_values(by='share', ascending=**False**) print("Sorted by Page total likes (descending):") sorted\_likes\_desc[['share']].head(30)

Sorted by Page total likes (descending):

[28]: share

333 181

3 147

|  |  |
| --- | --- |
| 105 | 139 |
| 252 | 128 |
| 244 | 123 |
| 222 | 122 |
| 264 | 109 |
| 353 | 99 |
| 159 | 98 |
| 315 | 97 |
| 101 | 95 |
| 139 | 90 |
| 99 | 90 |
| 219 | 90 |
| 71 | 84 |
| 361 | 80 |
| 335 | 78 |
| 90 | 77 |
| 243 | 76 |
| 75 | 74 |
| 26 | 72 |
| 327 | 70 |
| 67 | 70 |
| 28 | 63 |
| 98 | 61 |
| 342 | 61 |
| 140 | 60 |
| 201 | 58 |
| 212 | 58 |
| 82 | 57 |

# Transposing Data

[31]:

*# 1. Create a small subset for demonstration* small\_subset = fb\_data.iloc[:5, :5] print("Original data shape:", small\_subset.shape) display(small\_subset)

[33]:

*# 2. Transpose the data (rows become columns and vice versa)*

transposed\_data = small\_subset.T

print("Transposed data shape:", transposed\_data.shape)

Original data shape: (5, 5)

pagetotallikes type category postmonth postweekday

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 0 | 139441 | 2 | 2 | 12 | 4 |
| 1 | 139441 | 3 | 2 | 12 | 3 |
| 2 | 139441 | 2 | 3 | 12 | 3 |
| 3 | 139441 | 2 | 2 | 12 | 2 |
| 4 | 139441 | 2 | 2 | 12 | 2 |

display(transposed\_data)

Transposed data shape: (5, 5)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 |
| pagetotallikes | 139441 | 139441 | 139441 | 139441 | 139441 |
| type | 2 | 3 | 2 | 2 | 2 |
| category | 2 | 2 | 3 | 2 | 2 |
| postmonth | 12 | 12 | 12 | 12 | 12 |
| postweekday | 4 | 3 | 3 | 2 | 2 |

# Shape and Reshape Data

[36]:

*# 1. Check the current shape of the dataset*

print("Original dataset shape:", fb\_data.shape)

*# 2. Extract numeric columns for reshaping operations*

numeric\_cols = fb\_data.select\_dtypes(include=['float64', 'int64']).columns numeric\_data = fb\_data[numeric\_cols]

print("Numeric data shape:", numeric\_data.shape) print("Numeric columns:", list(numeric\_cols)[:5], "...")

[38]:

Original dataset shape: (372, 19)

Numeric data shape: (372, 19)

Numeric columns: ['pagetotallikes', 'type', 'category', 'postmonth', 'postweekday'] …

*# 3. Convert the dataframe to numpy array for reshaping* array\_data = numeric\_data.iloc[:100, :10].values print("Original array shape:", array\_data.shape)

Original array shape: (100, 10)

[40]:

*# 4. Reshape to 1D array*

reshaped\_1d = array\_data.reshape(-1) print("1D array shape:", reshaped\_1d.shape) print("First 10 elements:", reshaped\_1d[:10])

[42]:

1D array shape: (1000,)

First 10 elements: [139441 2 2 12 4 3 0 2752

5091 178]

*# 5. Reshape to different 2D shapes*

*# Calculate the total number of elements* total\_elements = array\_data.size print(f"Total elements: **{**total\_elements**}**")

*# Reshape to (20, total\_elements/20)*

cols = total\_elements // 20

reshaped\_2d = array\_data.reshape(20, cols)

print("Reshaped to 2D array with shape:", reshaped\_2d.shape) print("Sample of reshaped data:")

print(reshaped\_2d[:3, :5])

Total elements: 1000

Reshaped to 2D array with shape: (20, 50) Sample of reshaped data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | [[139441 | 2 | 2 | 12 | 4] |
| [139441 | 3 | 2 | 12 | 1] |
| [139441 | 3 | 2 | 12 | 5]] |
| [ ]: |  |  |  |  |  |
|  |  |  |  |  |  |
| [ ]: |  |  |  |  |  |