The data is available as a csv file, downland from kaggle. We will analyze this data using the pandas data frame import pandas as pd In [1]: data=pd.read_csv(r'C:\DA project\Covid-19 project\covid_19_india.csv') data Sno Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Cured Deaths Confirmed Out[3]: Date 1 30/01/20 6:00 PM Kerala 2 31/01/20 6:00 PM Kerala 0 0 3 01/02/20 6:00 PM 2 Kerala 0 0 2 4 02/02/20 6:00 PM 3 0 3 Kerala 0 0 5 03/02/20 6:00 PM 3 Kerala 0 0 3 **7781** 7782 27/10/20 8:00 AM Telengana - 213466 1315 232671 7782 7783 27/10/20 8:00 AM Tripura 28153 344 30293 7783 7784 27/10/20 8:00 AM Uttarakhand 55663 1001 60744 7784 7785 27/10/20 8:00 AM Uttar Pradesh - 438521 6904 472077 **7785** 7786 27/10/20 8:00 AM West Bengal - 310086 6546 353822 7786 rows × 9 columns data.count() 7786 Sno Out[4]: Date 7786 7786 Time State/UnionTerritory 7786 ConfirmedIndianNational 7786 ConfirmedForeignNational 7786 Cured 7786 Deaths 7786 Confirmed 7786 dtype: int64 data.isnull Time State/UnionTerritory ConfirmedIndianNational \ <bound method DataFrame.isnull of</pre> Sno Date Out[5]: 30/01/20 6:00 PM 1 Kerala 1 31/01/20 1 6:00 PM Kerala 1 01/02/20 2 3 6:00 PM Kerala 02/02/20 6:00 PM Kerala 3 3 5 03/02/20 6:00 PM Kerala 27/10/20 7781 7782 8:00 AM Telengana 7782 7783 27/10/20 8:00 AM Tripura 7783 7784 27/10/20 8:00 AM Uttarakhand 7784 7785 27/10/20 8:00 AM Uttar Pradesh 7785 7786 27/10/20 8:00 AM West Bengal ConfirmedForeignNational Cured Deaths Confirmed 0 0 0 0 1 0 1 0 0 1 2 0 0 0 2 3 0 0 0 3 4 0 0 0 3 213466 1315 232671 7781 7782 28153 344 30293 1001 7783 55663 60744 7784 438521 6904 472077 7785 310086 6546 353822 [7786 rows x 9 columns]> import seaborn as sns In [7] import matplotlib.pyplot as plt sns.heatmap(data.isnull()) In [8]: plt.show() 0.100 371 742 1113 1484 1855 2226 2597 2968 3339 4081 4452 4823 5194 5565 6307 6678 7049 7420 0.075 0.050 0.025 0.000 -0.025-0.050 -0.075 data.head(2) Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Cured Deaths Confirmed Out[9]: Date 1 30/01/20 6:00 PM 1 Kerala 2 31/01/20 6:00 PM Kerala data In [10]: ConfirmedIndianNational ConfirmedForeignNational Deaths Confirmed Out[10]: Sno Date Time State/UnionTerritory Cured 1 30/01/20 6:00 PM Kerala 2 31/01/20 6:00 PM Kerala 3 01/02/20 6:00 PM Kerala 2 0 4 02/02/20 6:00 PM Kerala 3 03/02/20 6:00 PM Kerala 0 0 Telengana **7781** 7782 27/10/20 8:00 AM - 213466 1315 232671 7782 7783 27/10/20 8:00 AM Tripura 28153 344 30293 7783 7784 27/10/20 8:00 AM Uttarakhand 55663 1001 60744 7784 7785 27/10/20 8:00 AM Uttar Pradesh 6904 472077 438521 7785 7786 27/10/20 8:00 AM West Bengal 310086 6546 353822 7786 rows × 9 columns Show the no. of Confirmed, Deaths and cured cases in each State/UnionTerritory data.head(2) Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Cured Deaths Confirmed Out[11]: 1 30/01/20 6:00 PM 1 Kerala 2 31/01/20 6:00 PM Kerala In [12]: data.groupby('State/UnionTerritory').sum() Deaths Confirmed Out[12]: Cured State/UnionTerritory **Andaman and Nicobar Islands** 861183 243399 3820 287887 Andhra Pradesh 864362 40409530 420087 48200952 637050 Arunachal Pradesh 857890 463296 1238 Assam 859191 10149044 44342 12661117 Bihar 863242 12246639 70787 14139595 Cases being reassigned to states 191339 0 0 345565 Chandigarh 864256 555075 9796 711347 Chhattisgarh 864369 4856690 59051 6790022 Dadar Nagar Haveli 83440 0 186 Dadra and Nagar Haveli and Daman and Diu 739184 212004 210 242811 0 2 Daman & Diu 2891 0 Delhi 866446 21499809 590548 25301742 **Goa** 863055 1831598 26885 2265168 Gujarat 864994 10033980 409257 12505593 Haryana 867042 7815830 103625 9269153 982711 Himachal Pradesh 865204 733347 11008 Jammu and Kashmir 867357 4073121 89189 5417186 Jharkhand 861467 4278351 47725 5346147 Karnataka 867667 32201170 649096 42342521 13533997 **Kerala** 868735 9290177 49465 **Ladakh** 868326 264244 3619 357010 Madhya Pradesh 866509 7540550 204437 9408126 Maharashtra 868554 78224180 3138377 105538919 Manipur 866105 601176 4634 833255 Meghalaya 853441 241289 3011 384180 Mizoram 866191 99350 137892 0 Nagaland 795331 342172 1239 480615 Odisha 868753 11773279 61090 14164794 Puducherry 868624 1357164 31549 1776066 **Punjab** 870033 213364 5818973 7305871 Rajasthan 870558 8862286 139444 10867315 **Sikkim** 791237 155187 2066 204479 Tamil Nadu 870770 42796330 797359 50167412 Telangana 168397 681436 13342 1077182 Telangana*** 40334 455 52466 13781886 Telengana 693649 11315346 88665 Telengana*** 4632 42909 480 57142 **Tripura** 861141 1287851 16686 1709637 Unassigned 161 **Uttar Pradesh** 871808 22563041 436558 28088092 Uttarakhand 870667 2389377 41841 3092164 403351 **West Bengal** 870576 16185659 19567105 data.groupby('State/UnionTerritory')['Cured'].sum().sort_values(ascending=False) State/UnionTerritory Out[13]: Maharashtra 78224180 Tamil Nadu 42796330 Andhra Pradesh 40409530 32201170 Karnataka Uttar Pradesh 22563041 Delhi 21499809 West Bengal 16185659 Bihar 12246639 **Odisha** 11773279 11315346 Telengana Assam 10149044 Gujarat 10033980 9290177 Kerala Rajasthan 8862286 Haryana 7815830 Madhya Pradesh 7540550 5818973 Punjab Chhattisgarh 4856690 Jharkhand 4278351 Jammu and Kashmir 4073121 Uttarakhand 2389377 1831598 Goa Puducherry 1357164 Tripura 1287851 Himachal Pradesh 733347 681436 Telangana Manipur 601176 Chandigarh 555075 Arunachal Pradesh 463296 342172 Nagaland Ladakh 264244 Andaman and Nicobar Islands 243399 241289 Meghalaya Dadra and Nagar Haveli and Daman and Diu 212004 Sikkim 155187 99350 Mizoram Telengana*** 42909 Telangana*** 40334 Dadar Nagar Haveli 14 Daman & Diu 0 Cases being reassigned to states 0 0 Unassigned Name: Cured, dtype: int64 data.groupby('State/UnionTerritory')['Cured'].sum().sort_values(ascending=False).head(20) #top 20 state Out[14]: State/UnionTerritory Maharashtra 78224180 Tamil Nadu 42796330 Andhra Pradesh 40409530 Karnataka 32201170 Uttar Pradesh 22563041 21499809 Delhi 16185659 West Bengal Bihar 12246639 11773279 0disha 11315346 Telengana 10149044 Assam Gujarat 10033980 9290177 Kerala Rajasthan 8862286 Haryana 7815830 Madhya Pradesh 7540550 5818973 Punjab Chhattisgarh 4856690 Jharkhand 4278351 Jammu and Kashmir 4073121 Name: Cured, dtype: int64 data.head(2) State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Out[15]: **Cured Deaths Confirmed** 1 30/01/20 6:00 PM Kerala 2 31/01/20 6:00 PM Kerala Remove all the records where cured is less than 1000 data.Cured < 1000 In [16]: 0 Out[16]: True True 2 True 3 True True 7781 False 7782 False 7783 False 7784 False 7785 False Name: Cured, Length: 7786, dtype: bool In [17] data[data.Cured<1000] Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Cured Deaths Confirmed Out[17]: Sno Date 1 30/01/20 6:00 PM 0 0 Kerala 0 0 1 2 31/01/20 6:00 PM Kerala 3 01/02/20 6:00 PM Kerala 2 0 0 0 2 4 02/02/20 6:00 PM Kerala 5 03/02/20 6:00 PM 3 0 0 0 3 Kerala 6338 6339 16/09/20 8:00 AM 1480 Mizoram 922 0 6373 6374 17/09/20 8:00 AM Mizoram 939 1506 **6408** 6409 18/09/20 8:00 AM 0 Mizoram 949 1534 6443 6444 19/09/20 8:00 AM Mizoram 973 1548 6478 6479 20/09/20 8:00 AM Mizoram 990 0 1578 3250 rows × 9 columns In which State/UnionTerritory, maximum no. of confirmed cases data.head(2) In [18]: Out[18]: Sno Date Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational **Cured Deaths Confirmed** 1 30/01/20 6:00 PM 0 0 Kerala 1 2 31/01/20 6:00 PM 1 Kerala data.groupby('State/UnionTerritory').Confirmed.sum() State/UnionTerritory Out[19]: Andaman and Nicobar Islands 287887 Andhra Pradesh 48200952 637050 Arunachal Pradesh Assam 12661117 14139595 Bihar Cases being reassigned to states 345565 711347 Chandigarh Chhattisgarh 6790022 Dadar Nagar Haveli 186 Dadra and Nagar Haveli and Daman and Diu 242811 Daman & Diu Delhi 25301742 Goa 2265168 Gujarat 12505593 9269153 Haryana Himachal Pradesh 982711 Jammu and Kashmir 5417186 Jharkhand 5346147 Karnataka 42342521 Kerala 13533997 357010 Ladakh 9408126 Madhya Pradesh Maharashtra 105538919 Manipur 833255 384180 Meghalaya Mizoram 137892 Nagaland 480615 14164794 0disha 1776066 Puducherry Punjab 7305871 Rajasthan 10867315 Sikkim 204479 Tamil Nadu 50167412 Telangana 1077182 Telangana * * * 52466 13781886 Telengana Telengana*** 57142 1709637 Tripura Unassigned 161 Uttar Pradesh 28088092 3092164 Uttarakhand West Bengal 19567105 Name: Confirmed, dtype: int64 In [20]: data.groupby('State/UnionTerritory').Confirmed.sum().sort_values(ascending=False)#maximum State/UnionTerritory Out[20]: Maharashtra 105538919 Tamil Nadu 50167412 Andhra Pradesh 48200952 Karnataka 42342521 28088092 Uttar Pradesh 25301742 Delhi West Bengal 19567105 0disha 14164794 Bihar 14139595 Telengana 13781886 13533997 Kerala 12661117 Assam Gujarat 12505593 10867315 Rajasthan Madhya Pradesh 9408126 Haryana 9269153 Punjab 7305871 Chhattisgarh 6790022 Jammu and Kashmir 5417186 Jharkhand 5346147 Uttarakhand 3092164 2265168 Goa Puducherry 1776066 Tripura 1709637 Telangana 1077182 Himachal Pradesh 982711 833255 Manipur Chandigarh 711347 Arunachal Pradesh 637050 Nagaland 480615 Meghalaya 384180 357010 Ladakh Cases being reassigned to states 345565 Andaman and Nicobar Islands 287887 Dadra and Nagar Haveli and Daman and Diu 242811 Sikkim 204479 Mizoram 137892 Telengana*** 57142 Telangana*** 52466 Dadar Nagar Haveli 186 Unassigned 161 Daman & Diu 2 Name: Confirmed, dtype: int64 data.groupby('State/UnionTerritory').Confirmed.sum().sort_values(ascending=False).head(10) #top 10 state State/UnionTerritory Out[21]: Maharashtra 105538919 Tamil Nadu 50167412 48200952 Andhra Pradesh 42342521 Karnataka Uttar Pradesh 28088092 Delhi 25301742 West Bengal 19567105 0disha 14164794 Bihar 14139595 Telengana 13781886 Name: Confirmed, dtype: int64 In which State/UnionTerritory, minimum no. of Deaths were recorded? data.head(2) In [22] ConfirmedForeignNational Cured Deaths Confirmed Sno Date Time State/UnionTerritory ConfirmedIndianNational Out[22]: 1 30/01/20 6:00 PM Kerala 0 0 1 2 31/01/20 6:00 PM 0 Kerala 0 1 data.groupby('State/UnionTerritory').Deaths.sum() State/UnionTerritory Out[23]: 3820 Andaman and Nicobar Islands Andhra Pradesh 420087 Arunachal Pradesh 1238 Assam 44342 Bihar 70787 Cases being reassigned to states Chandigarh 9796 Chhattisgarh 59051 Dadar Nagar Haveli 0 Dadra and Nagar Haveli and Daman and Diu 210 Daman & Diu 0 Delhi 590548 Goa 26885 Gujarat 409257 Haryana 103625 Himachal Pradesh 11008 Jammu and Kashmir 89189 47725 Jharkhand Karnataka 649096 Kerala 49465 Ladakh 3619 Madhya Pradesh 204437 3138377 Maharashtra Manipur 4634 Meghalaya 3011 Mizoram 0 Nagaland 1239 0disha 61090 Puducherry 31549 Punjab 213364 Rajasthan 139444 Sikkim 2066 Tamil Nadu 797359 Telangana 13342 Telangana * * * 455 Telengana 88665 Telengana*** 480 16686 Tripura Unassigned Uttar Pradesh 436558 Uttarakhand 41841 West Bengal 403351 Name: Deaths, dtype: int64 In [24]: data.groupby('State/UnionTerritory').Deaths.sum().sort_values(ascending=True) #min k liye acending true kr denge State/UnionTerritory Unassigned Cases being reassigned to states 0 0 Dadar Nagar Haveli Daman & Diu 0 Mizoram Dadra and Nagar Haveli and Daman and Diu 210 Telangana * * * 455 Telengana*** 480 Arunachal Pradesh 1238 Nagaland 1239 Sikkim 2066 Meghalaya 3011 Ladakh 3619 Andaman and Nicobar Islands 3820 4634 Manipur Chandigarh 9796 Himachal Pradesh 11008 Telangana 13342 16686 Tripura 26885 Goa Puducherry 31549 Uttarakhand 41841 44342 Assam Jharkhand 47725 49465 Kerala Chhattisgarh 59051 0disha 61090 Bihar 70787 88665 Telengana Jammu and Kashmir 89189 Haryana 103625 Rajasthan 139444 Madhya Pradesh 204437 Punjab 213364 West Bengal 403351 Gujarat 409257 Andhra Pradesh 420087 Uttar Pradesh 436558 Delhi 590548 649096 Karnataka Tamil Nadu 797359 Maharashtra 3138377 Name: Deaths, dtype: int64 In [26]: data.groupby('State/UnionTerritory').Deaths.sum().sort_values(ascending=True).head(10) #top 10 State/UnionTerritory State/UnionTerritory Out[26]: Unassigned 0 Cases being reassigned to states 0 Dadar Nagar Haveli Daman & Diu 0 Mizoram 0 Dadra and Nagar Haveli and Daman and Diu 210 Telangana** 455 Telengana*** 480 Arunachal Pradesh 1238 Nagaland 1239 Name: Deaths, dtype: int64 How many confirmed, deaths and cured were reported from Jammu and Kashmir data.head(2) In [27]: ConfirmedForeignNational Cured Deaths Confirmed Time State/UnionTerritory ConfirmedIndianNational Out[27]: Sno Date 1 30/01/20 6:00 PM 0 Kerala 0 0 1 1 2 31/01/20 6:00 PM Kerala 0 data['State/UnionTerritory']=='Jammu and Kashmir'# True false mein dega In [32]: False Out[32]: 0 False 2 False 3 False False 7781 False 7782 False 7783 False False 7785 False Name: State/UnionTerritory, Length: 7786, dtype: bool data[data['State/UnionTerritory']=='Jammu and Kashmir'] In [37]: Out[37]: Sno Date Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Cured Deaths Confirmed 82 09/03/20 6:00 PM 81 0 0 Jammu and Kashmir 1 97 10/03/20 6:00 PM Jammu and Kashmir 107 11/03/20 6:00 PM 0 0 106 1 0 1 Jammu and Kashmir 118 12/03/20 6:00 PM Jammu and Kashmir 131 13/03/20 6:00 PM Jammu and Kashmir 0 0 0 1 130 1 **7624** 7625 23/10/20 8:00 AM 1412 - 80802 90166 Jammu and Kashmir **7659** 7660 24/10/20 8:00 AM Jammu and Kashmir - 81486 1424 90752 **7694** 7695 25/10/20 8:00 AM Jammu and Kashmir - 82219 1430 91329 7730 26/10/20 8:00 AM Jammu and Kashmir - 82858 91861 **7764** 7765 27/10/20 8:00 AM - 83485 92225 Jammu and Kashmir 1444 233 rows × 9 columns Sort the entire data wrt no. of confirmed cases in ascending order data.head(2) In [38]: Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Cured Deaths Confirmed Out[38]: Date 1 30/01/20 6:00 PM Kerala 2 31/01/20 6:00 PM Kerala 0 data.sort_values(by=['Confirmed']) #bydefault ascending order mein show krega Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational **Cured Deaths Confirmed** Out[39]: Sno Date **1047** 1048 17/04/20 5:00 PM Nagaland 0 **1113** 1114 19/04/20 5:00 PM 0 0 Nagaland **1179** 1180 21/04/20 5:00 PM Nagaland 0 0 0 **1146** 1147 20/04/20 5:00 PM Nagaland 0 0 **1080** 1081 18/04/20 5:00 PM Nagaland 0 0 7630 7631 23/10/20 8:00 AM Maharashtra - 1431856 42831 1625197 **7665** 7666 24/10/20 8:00 AM - 1445103 43015 1632544 Maharashtra 7700 7701 25/10/20 8:00 AM - 1455107 43152 1638961 Maharashtra 7735 7736 26/10/20 8:00 AM Maharashtra - 1460755 43264 1645020 - 1470660 **7770** 7771 27/10/20 8:00 AM Maharashtra 43348 1648665 7786 rows × 9 columns data.sort_values(by=['Confirmed']).head(50) Time State/UnionTerritory ConfirmedIndianNational ConfirmedForeignNational Cured Deaths Confirmed Out[40]: Date **1047** 1048 17/04/20 5:00 PM Nagaland 0 0 **1113** 1114 19/04/20 5:00 PM Nagaland **1179** 1180 21/04/20 5:00 PM 0 0 0 Nagaland **1146** 1147 20/04/20 5:00 PM Nagaland **1080** 1081 18/04/20 5:00 PM 0 Nagaland 0 0 982 15/04/20 5:00 PM 981 Nagaland 0 **1014** 1015 16/04/20 5:00 PM Nagaland 0 0 0 Arunachal Pradesh 2273 2274 25/05/20 8:00 AM 0 0 2263 2264 24/05/20 8:00 AM Sikkim 1 0 2258 2259 24/05/20 8:00 AM Mizoram 2238 2239 24/05/20 8:00 AM Arunachal Pradesh 1 0 1 577 02/04/20 6:00 PM 0 0 576 Manipur 578 02/04/20 6:00 PM 0 577 0 1 Mizoram 0 2224 2225 23/05/20 8:00 AM Mizoram 0 2209 2210 23/05/20 8:00 AM Dadar Nagar Haveli 0 1 570 571 02/04/20 6:00 PM Jharkhand 2298 2299 25/05/20 8:00 AM Sikkim 0 0 1 1 30/01/20 6:00 PM 0 0 Kerala 547 548 01/04/20 7:30 PM 0 0 Manipur 1 542 01/04/20 7:30 PM 0 Jharkhand 0 0 2328 2329 26/05/20 8:00 AM Mizoram 1 1 0 2334 2335 26/05/20 8:00 AM Sikkim 0 0 0 531 532 01/04/20 7:30 PM Assam 1 2364 2365 27/05/20 8:00 AM Mizoram 0 0 2370 2371 27/05/20 8:00 AM Sikkim 1 520 521 31/03/20 8:30 PM Puducherry 0 0 519 31/03/20 8:30 PM 0 0 518 Mizoram 1 518 31/03/20 8:30 PM 0 0 517 Manipur 2400 2401 28/05/20 8:00 AM 0 Mizoram 1 1 549 01/04/20 7:30 PM 0 0 548 Mizoram 0 2204 2205 23/05/20 8:00 AM Arunachal Pradesh 1 1 **2175** 2176 22/05/20 8:00 AM Dadar Nagar Haveli 2190 2191 22/05/20 8:00 AM 1 0 Mizoram 1 0 0 2008 2009 17/05/20 8:00 AM Dadar Nagar Haveli 669 05/04/20 6:00 PM 668 0 0 Mizoram 1 2023 2024 17/05/20 8:00 AM Mizoram 0 Arunachal Pradesh 2036 2037 18/05/20 8:00 AM 1 0 1 Dadar Nagar Haveli 0 **2041** 2042 18/05/20 8:00 AM 0 Arunachal Pradesh 0 650 651 05/04/20 6:00 PM 0 1 2056 2057 18/05/20 8:00 AM 0 Mizoram 0 2069 2070 19/05/20 8:00 AM Arunachal Pradesh 1 1 0 0 2074 2075 19/05/20 8:00 AM Dadar Nagar Haveli 639 04/04/20 6:00 PM 0 0 638 Mizoram 1 2090 19/05/20 8:00 AM 2089 Mizoram Arunachal Pradesh 0 2102 2103 20/05/20 8:00 AM 1 1 **2107** 2108 20/05/20 8:00 AM Dadar Nagar Haveli 0 0 621 04/04/20 6:00 PM Arunachal Pradesh 0 0 620 1 2122 2123 20/05/20 8:00 AM Mizoram **2136** 2137 21/05/20 8:00 AM Arunachal Pradesh **2141** 2142 21/05/20 8:00 AM Dadar Nagar Haveli