Exploratory_Data_Analysis_IndianCities

October 8, 2020

1 Exploratory Data Analysis on Indian states

Click here for final insights

I am exploring this data set for any insights. I am going to plot bar graphs and actual India map against various attributes. Lets explore India :)

This is going to be a long notebook. So if you want to check out the results quickly, follow the below links :

Which State have most number of top cities ?

Top 10 Populous cities in India Which state have most of its population in urban areas ?

Which state have most of its male population in urban areas ? Top 10 cities with high male population

Which state have most of its male population in urban areas ? Top 10 cities with high female population

Which state have most of its kids population in urban areas ? Top 10 cities with high kids population

Which state have most of its male kids population in urban areas ? Top 10 cities with high male kids population

Which state have most of its female kids population in urban areas ? Top 10 cities with high female kids population

Analysing Litracy rate of the states Top 10 cities with most number of literates live

Analysing Male Litracy rate of the states Top 10 cities with most number of male literates live

Analysing Female Litracy rate of the states Top 10 cities with most number of female literates live

Analyzing effective literacy rate

Analyzing Graduates

Analyzing Sex ratio

Analyzing Sex ratio for children below 6

1.1 Importing all the required packages

```
In [1]: # importing packages
        import pandas as pd
        import numpy as np
        from scipy.interpolate import spline
        from numpy import array
        import matplotlib as mpl
```

```
# for plots
import matplotlib.pyplot as plt
from matplotlib import cm
from matplotlib.dates import date2num
from mpl_toolkits.basemap import Basemap

# for date and time processing
import datetime

# for statistical graphs
import seaborn as sns
```

1.2 Importing all the Data into notebook

```
In [2]: cities = pd.read_csv ("./Indian_cities.csv")
```

1.3 Viewing Data and Verifying

```
In [3]: cities.head ()
```

Out[3]:

	name_of_city	state_code	state_name	dist_code	population_total	population_male	population_female	0-6_population_total	0-6_population_male	0-6_population_female	...	literates_female	sex_ratio	child_sex_ratio	effective_literacy_rate_total
0	Abohar	3	PUNJAB	9	145238	76840	68398	15870	8587	7283	...	44972	890	848	79.86
1	Achalpur	27	MAHARASHTRA	7	112293	58256	54037	11810	6186	5624	...	43086	928	909	91.99
2	Adilabad	28	ANDHRA PRADESH	1	117388	59232	58156	13103	6731	6372	...	37660	982	947	80.51
3	Adityapur	20	JHARKHAND	24	173988	91495	82493	23042	12063	10979	...	54515	902	910	83.46
4	Adoni	28	ANDHRA PRADESH	21	166537	82743	83794	18406	9355	9051	...	45089	1013	968	68.38

```
   effective_literacy_rate_male  effective_literacy_rate_female  \
0                         85.49                           73.59
1                         94.77                           89.00
2                         88.18                           72.73
3                         89.98                           76.23
4                         76.58                           60.33

                location  total_graduates male_graduates  female_graduates
0  30.1452928,74.1993043            16287           8612              7675
1   21.257584,77.5086754             8863           5269              3594
2   19.0809075,79.560344            10565           6797              3768
3  22.7834741,86.1576889            19225          12189              7036
4  15.6322227,77.2728368            11902           7871              4031

[5 rows x 22 columns]
```

In [4]:
```
cities.info ()
# there is no null values anywhere in the dataset
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 493 entries, 0 to 492
Data columns (total 22 columns):
name_of_city                    493 non-null object
state_code                      493 non-null int64
state_name                      493 non-null object
dist_code                       493 non-null int64
population_total                493 non-null int64
population_male                 493 non-null int64
population_female               493 non-null int64
0-6_population_total            493 non-null int64
0-6_population_male             493 non-null int64
0-6_population_female           493 non-null int64
literates_total                 493 non-null int64
literates_male                  493 non-null int64
literates_female                493 non-null int64
sex_ratio                       493 non-null int64
child_sex_ratio                 493 non-null int64
effective_literacy_rate_total   493 non-null float64
effective_literacy_rate_male    493 non-null float64
effective_literacy_rate_female  493 non-null float64
location                        493 non-null object
total_graduates                 493 non-null int64
male_graduates                  493 non-null int64
female_graduates                493 non-null int64
dtypes: float64(3), int64(16), object(3)
memory usage: 84.8+ KB
```

In [5]:
```
cities.describe ()
```

Out[5]:

	state_code	dist_code	population_total	population_male
count	493.000000	493.000000	4.930000e+02	4.930000e+02
mean	18.643002	16.782961	4.481124e+05	2.343468e+05
std	9.297168	15.566131	1.033228e+06	5.487786e+05
min	1.000000	1.000000	1.000360e+05	5.020100e+04
25%	9.000000	7.000000	1.261420e+05	6.638400e+04
50%	19.000000	13.000000	1.841330e+05	9.665500e+04
75%	27.000000	21.000000	3.490330e+05	1.750550e+05
max	35.000000	99.000000	1.247845e+07	6.736815e+06

	population_female	0-6_population_total	0-6_population_male
count	4.930000e+02	4.930000e+02	493.000000
mean	2.137656e+05	4.709285e+04	24849.527383
std	4.848622e+05	1.050279e+05	55535.310272
min	4.512600e+04	6.547000e+03	3406.000000
25%	6.041100e+04	1.363900e+04	7221.000000
50%	8.776800e+04	1.944000e+04	10342.000000
75%	1.700260e+05	3.794500e+04	19982.000000
max	5.741632e+06	1.209275e+06	647938.000000

	0-6_population_female	literates_total	literates_male
count	493.000000	4.930000e+02	4.930000e+02
mean	22243.320487	3.461527e+05	1.894384e+05
std	49523.241379	8.220952e+05	4.534753e+05
min	3107.000000	5.699800e+04	3.475100e+04
25%	6457.000000	9.768700e+04	5.357800e+04
50%	9172.000000	1.413290e+05	7.590600e+04
75%	17954.000000	2.679000e+05	1.455480e+05
max	561337.000000	1.023759e+07	5.727774e+06

	literates_female	sex_ratio	child_sex_ratio
count	4.930000e+02	493.000000	493.000000
mean	1.567143e+05	930.294118	902.332657
std	3.690677e+05	55.849106	49.794689
min	2.224700e+04	700.000000	762.000000
25%	4.391400e+04	890.000000	868.000000
50%	6.383600e+04	922.000000	903.000000
75%	1.235030e+05	971.000000	942.000000
max	4.509812e+06	1093.000000	1185.000000

	effective_literacy_rate_total	effective_literacy_rate_male
count	493.000000	493.000000
mean	85.131460	89.920162
std	6.186345	5.377492
min	49.510000	52.270000
25%	81.750000	87.280000
50%	85.970000	91.180000
75%	89.330000	93.400000

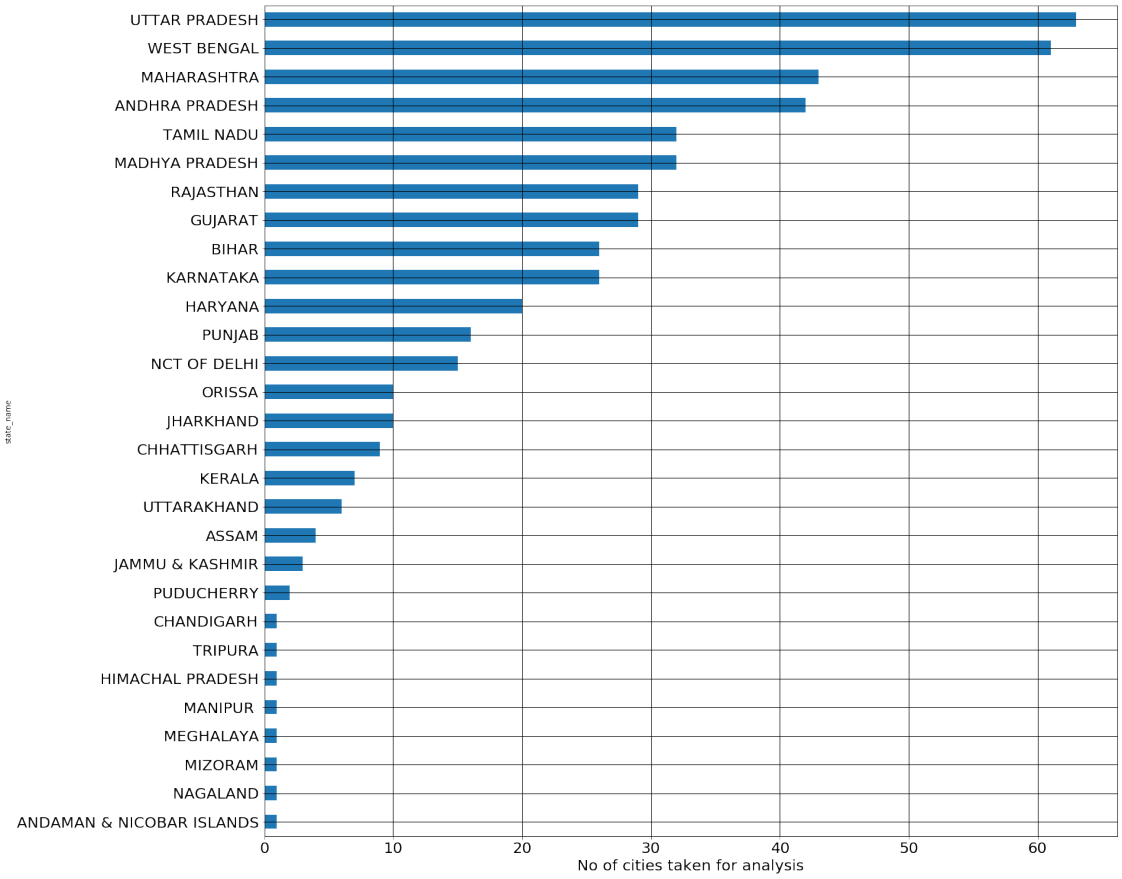
```
max                          98.800000                    99.300000

       effective_literacy_rate_female  total_graduates  male_graduates  \
count                      493.000000     4.930000e+02    4.930000e+02
mean                        79.967181     6.620236e+04    3.771556e+04
std                          7.577825     1.778187e+05    9.849574e+04
min                         46.450000     2.532000e+03    1.703000e+03
25%                         75.800000     1.527700e+04    9.289000e+03
50%                         80.920000     2.395900e+04    1.404900e+04
75%                         85.400000     5.036700e+04    2.787200e+04
max                         98.310000     2.221137e+06    1.210040e+06

       female_graduates
count      4.930000e+02
mean       2.848680e+04
std        7.951556e+04
min        8.290000e+02
25%        6.114000e+03
50%        9.558000e+03
75%        2.086600e+04
max        1.011097e+06
```

```
In [6]: print (cities.describe(include=['O']))
        # from the below output, we can learn that there is two Aurangabad's. One is in Mahara
        # and one is in Bihar
        # most of the cities are selected from Uttar Pradesh
```

```
       name_of_city     state_name               location
count           493            493                    493
unique          492             29                    490
top     Aurangabad   UTTAR PRADESH  22.6500518,88.3423199
freq              2             63                      2
```

1.4 Plotting state wise cities to check which state have most number of cities in it

```
In [7]: # A bar chart to show from which states, how many cities are taken for examination.
        fig = plt.figure(figsize=(20,20))
        states = cities.groupby('state_name')['name_of_city'].count().sort_values(ascending=Tru
        states.plot(kind="barh", fontsize = 20)
        plt.grid(b=True, which='both', color='Black',linestyle='-')
        plt.xlabel('No of cities taken for analysis', fontsize = 20)
        plt.show ()
        # we can see states like UP and WB are given high priority by taking more than 60 citi
```



1.5 Top 10 populous cities

```
In [8]: # Extracting Co-ordinates details from the provided data
        cities['latitude'] = cities['location'].apply(lambda x: x.split(',')[0])
        cities['longitude'] = cities['location'].apply(lambda x: x.split(',')[1])
        cities.head(1)
```

```
Out[8]:   name_of_city  state_code state_name  dist_code  population_total  \
        0      Abohar            3     PUNJAB          9            145238

           population_male  population_female  0-6_population_total  \
        0            76840              68398                 15870

           0-6_population_male  0-6_population_female  ...  child_sex_ratio  \
        0                 8587                   7283  ...              848

           effective_literacy_rate_total  effective_literacy_rate_male  \
        0                          79.86                         85.49

           effective_literacy_rate_female              location  total_graduates  \
```

```
0                          73.59  30.1452928,74.1993043            16287

   male_graduates  female_graduates    latitude   longitude
0            8612              7675  30.1452928  74.1993043

[1 rows x 24 columns]
```

In [9]:
```
# A table to show top 10 cities with most population
print("The Top 10 Cities sorted according to the Total Population (Descending Order)")
top_pop_cities = cities.sort_values(by='population_total',ascending=False)
top10_pop_cities=top_pop_cities.head(10)
top10_pop_cities
```

```
The Top 10 Cities sorted according to the Total Population (Descending Order)
```

Out[9]:

```
           name_of_city  state_code      state_name  dist_code  \
185     Greater Mumbai           27     MAHARASHTRA         99
141              Delhi            7    NCT OF DELHI         99
72            Bengaluru          29       KARNATAKA         18
184  Greater Hyderabad           28  ANDHRA PRADESH         99
7             Ahmadabad          24         GUJARAT          7
119             Chennai          33      TAMIL NADU          2
274             Kolkata          19     WEST BENGAL         16
449               Surat          24         GUJARAT         25
380                Pune          27     MAHARASHTRA         25
225              Jaipur           8       RAJASTHAN         12

     population_total  population_male  population_female  \
185          12478447          6736815            5741632
141          11007835          5871362            5136473
72            8425970          4401299            4024671
184           6809970          3500802            3309168
7             5570585          2935869            2634716
119           4681087          2357633            2323454
274           4486679          2362662            2124017
449           4462002          2538243            1923759
380           3115431          1602137            1513294
225           3073350          1619280            1454070

     0-6_population_total  0-6_population_male  0-6_population_female  ...  \
185                1139146               599007                 540139  ...
141                1209275               647938                 561337  ...
72                  862493               444639                 417854  ...
184                 725816               373794                 352022  ...
7                   589076               317917                 271159  ...
119                 418541               213084                 205457  ...
274                 300052               155475                 144577  ...
```

```
449                531522                293208                238314  ...
380                324572                171152                153420  ...
225                378788                204320                174468  ...

     child_sex_ratio  effective_literacy_rate_total  \
185              902                          90.28
141              866                          87.60
72               940                          89.59
184              942                          82.96
7                853                          89.62
119              964                          90.33
274              930                          87.14
449              813                          89.03
380              896                          91.61
225              854                          84.34

     effective_literacy_rate_male  effective_literacy_rate_female  \
185                         93.32                           86.70
141                         91.44                           83.20
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
119                         93.47                           87.16
274                         89.08                           84.98
449                         92.76                           84.05
380                         95.13                           87.91
225                         90.61                           77.41

                location  total_graduates  male_graduates  female_graduates  \
185       19.0760,72.8777          1802371          964964            837407
141       28.7041,77.1025          2221137         1210040           1011097
72        12.9716,77.5946          1591163          908363            682800
184       17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
119  13.0826802,80.2707184           879695          487428            392267
274    22.572646,88.363895           818476          461615            356861
449  21.1702401,72.8310607           278795          160566            118229
380  18.5204303,73.8567437           656508          349022            307486
225  26.9124336,75.7872709           533148          319107            214041

       latitude   longitude
185     19.0760     72.8777
141     28.7041     77.1025
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
119  13.0826802  80.2707184
274   22.572646   88.363895
```

```
449  21.1702401  72.8310607
380  18.5204303  73.8567437
225  26.9124336  75.7872709

[10 rows x 24 columns]
```

```
In [10]: # Plotting these top 10 populous cities on India map. Circles are sized according to
         # population of the respective city

         plt.subplots(figsize=(20, 15))
         map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                       llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

         map.drawmapboundary ()
         map.drawcountries ()
         map.drawcoastlines ()

         lg=array(top10_pop_cities['longitude'])
         lt=array(top10_pop_cities['latitude'])
         pt=array(top10_pop_cities['population_total'])
         nc=array(top10_pop_cities['name_of_city'])

         x, y = map(lg, lt)
         population_sizes = top10_pop_cities["population_total"].apply(lambda x: int(x / 5000))
         plt.scatter(x, y, s=population_sizes, marker="o", c=population_sizes, cmap=cm.Dark2, a


         for ncs, xpt, ypt in zip(nc, x, y):
             plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

         plt.title('Top 10 Populated Cities in India',fontsize=20)
```

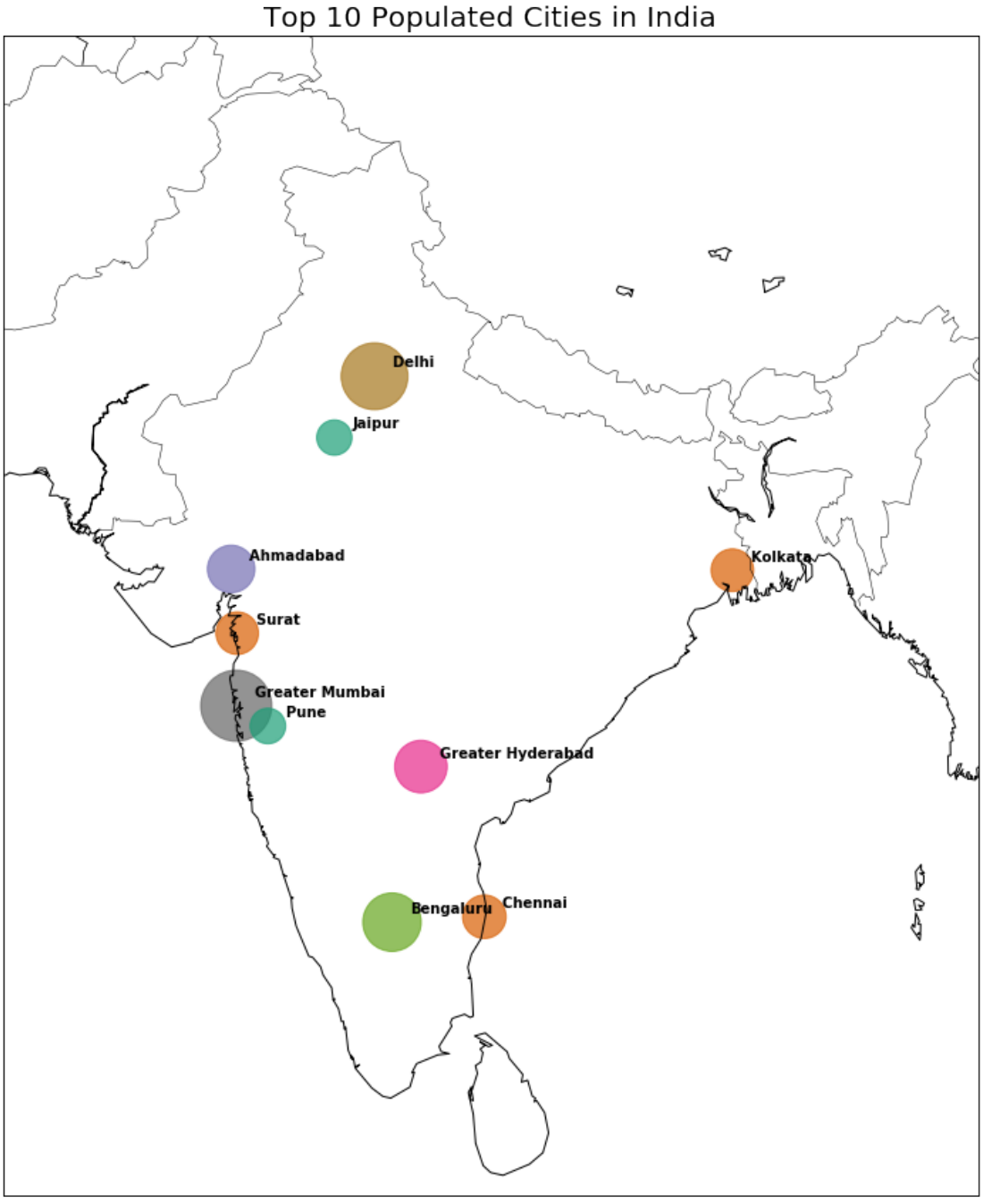
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

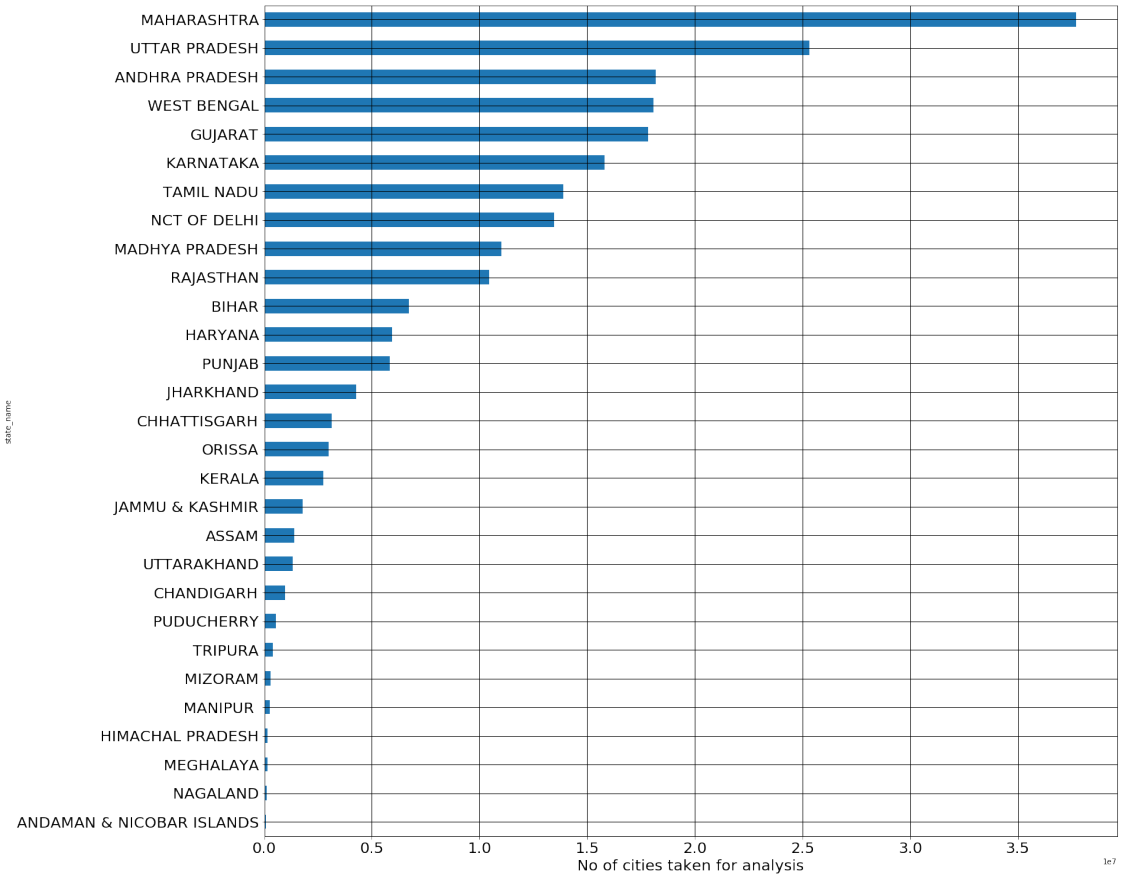
```
Out[10]: Text(0.5,1,'Top 10 Populated Cities in India')
```

1.6 Plotting Statewise cities to check which state have most population living in urban areas

```
In [11]: # A bar chart to show the population of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['population_total'].sum().sort_values(ascending=
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of cities taken for analysis', fontsize = 20)
```

```
plt.show ()
# we can see states like Maharashtra and UP have huge urban population
```





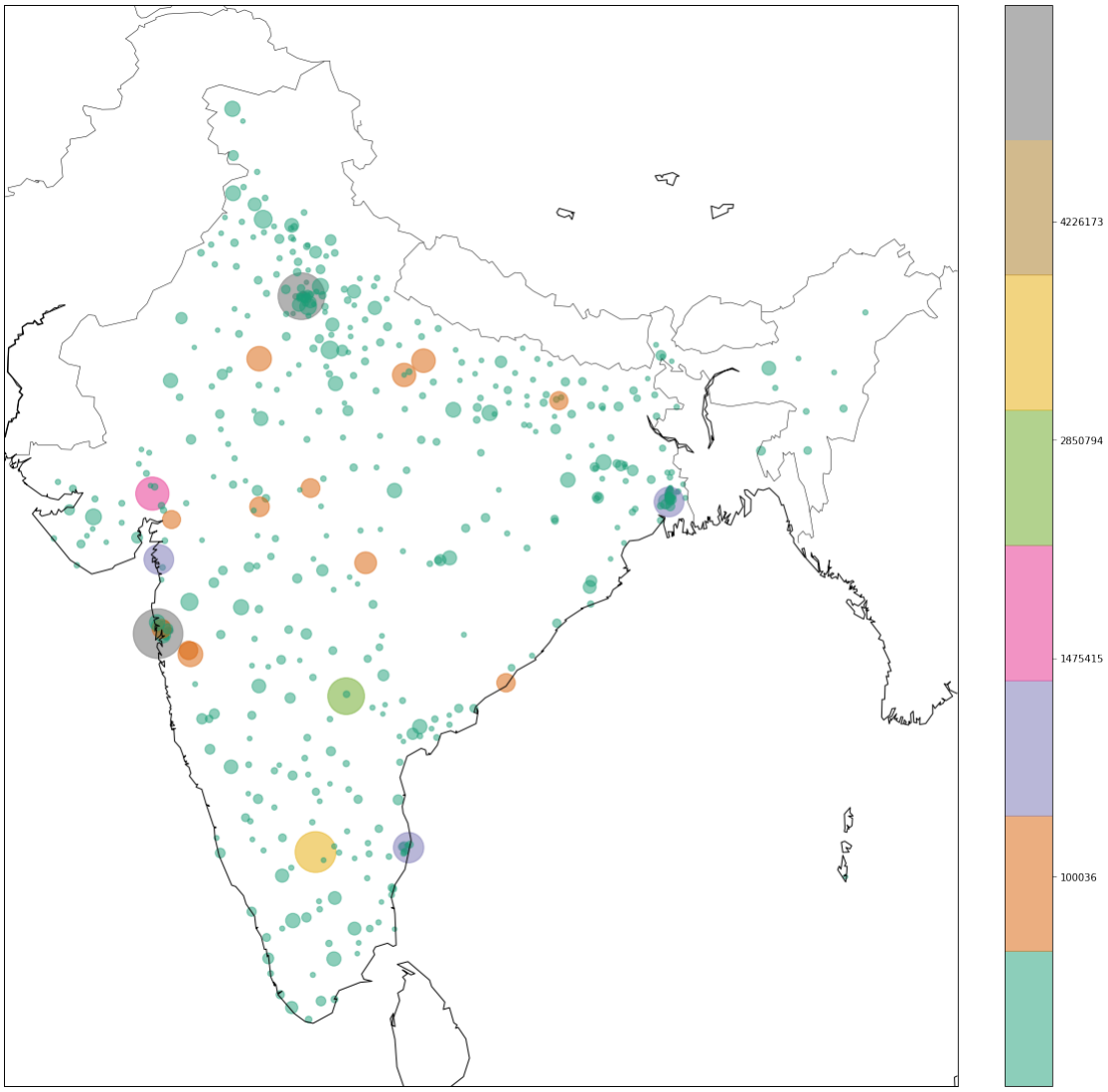
1.7 Plotting every city on India map according to population

In [12]:
```
# Creating a function to plot the population data on real India map

def plot_map(sizes, colorbarValue):

    plt.figure(figsize=(19,20))
    f, ax = plt.subplots(figsize=(19, 20))

    # Setting up Basemap
    map = Basemap(width=5000000, height=3500000, resolution='l', projection='aea', llc
                  llcrnrlat=6, urcrnrlon=99, urcrnrlat=36, lon_0=78, lat_0=20, ax=ax)

    # draw map boundaries
    map.drawmapboundary()
    map.drawcountries()
    map.drawcoastlines()

    # plotting cities on map using previously derived coordinates
    x, y = map(array(cities["longitude"]), array(cities["latitude"]))
```

```
        cs = map.scatter(x, y, s=sizes, marker="o", c=sizes, cmap=cm.Dark2, alpha=0.5)

        # adding colorbar
        cbar = map.colorbar(cs, location='right',pad="5%")
        cbar.ax.set_yticklabels(colorbarValue)

        plt.show()
```

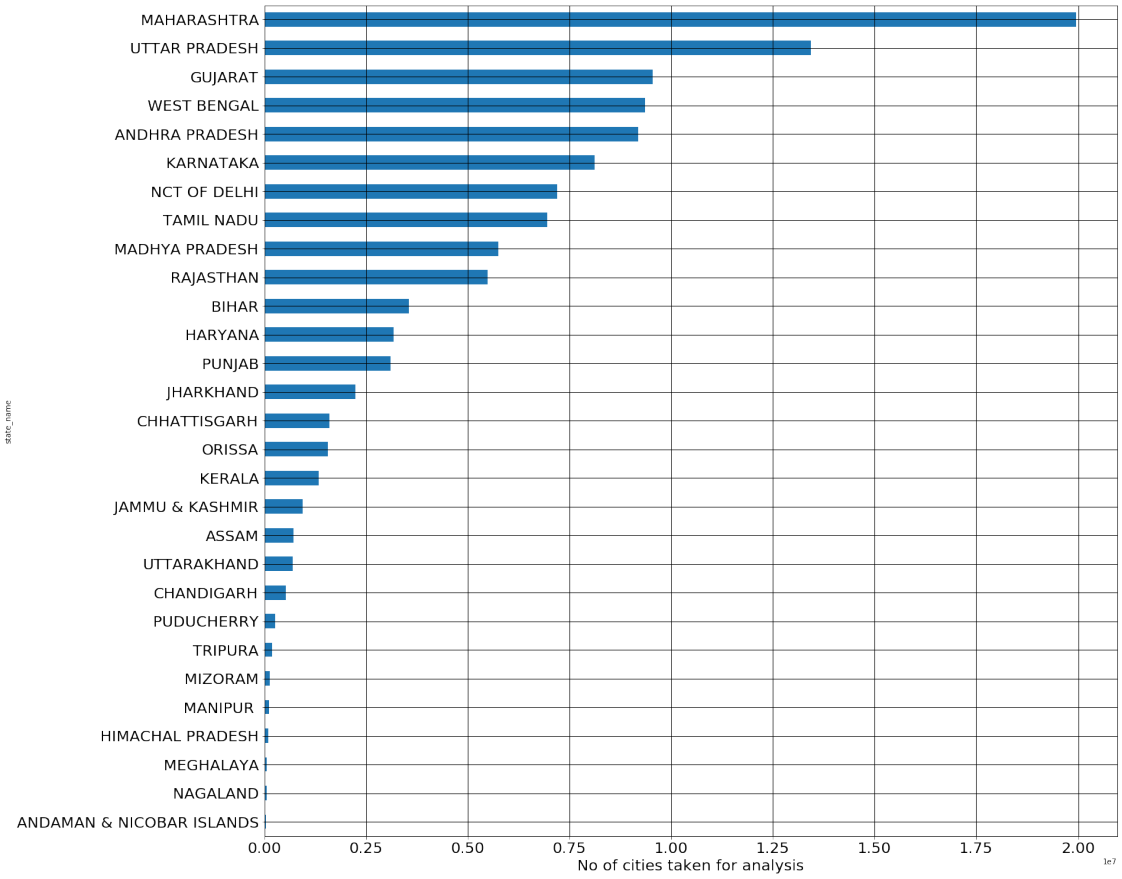
```
In [13]: # Using the function created in the previous cell, we are plotting the population dat

        population_sizes = cities["population_total"].apply(lambda x: int(x / 5000))
        colorbarValue = np.linspace(cities["population_total"].min(), cities["population_total
                                    num=10)
        colorbarValue = colorbarValue.astype(int)

        plot_map(population_sizes, colorbarValue)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)


<matplotlib.figure.Figure at 0x25e631e3860>
```



1.8 Plotting Statewise cities to check which state have most male population

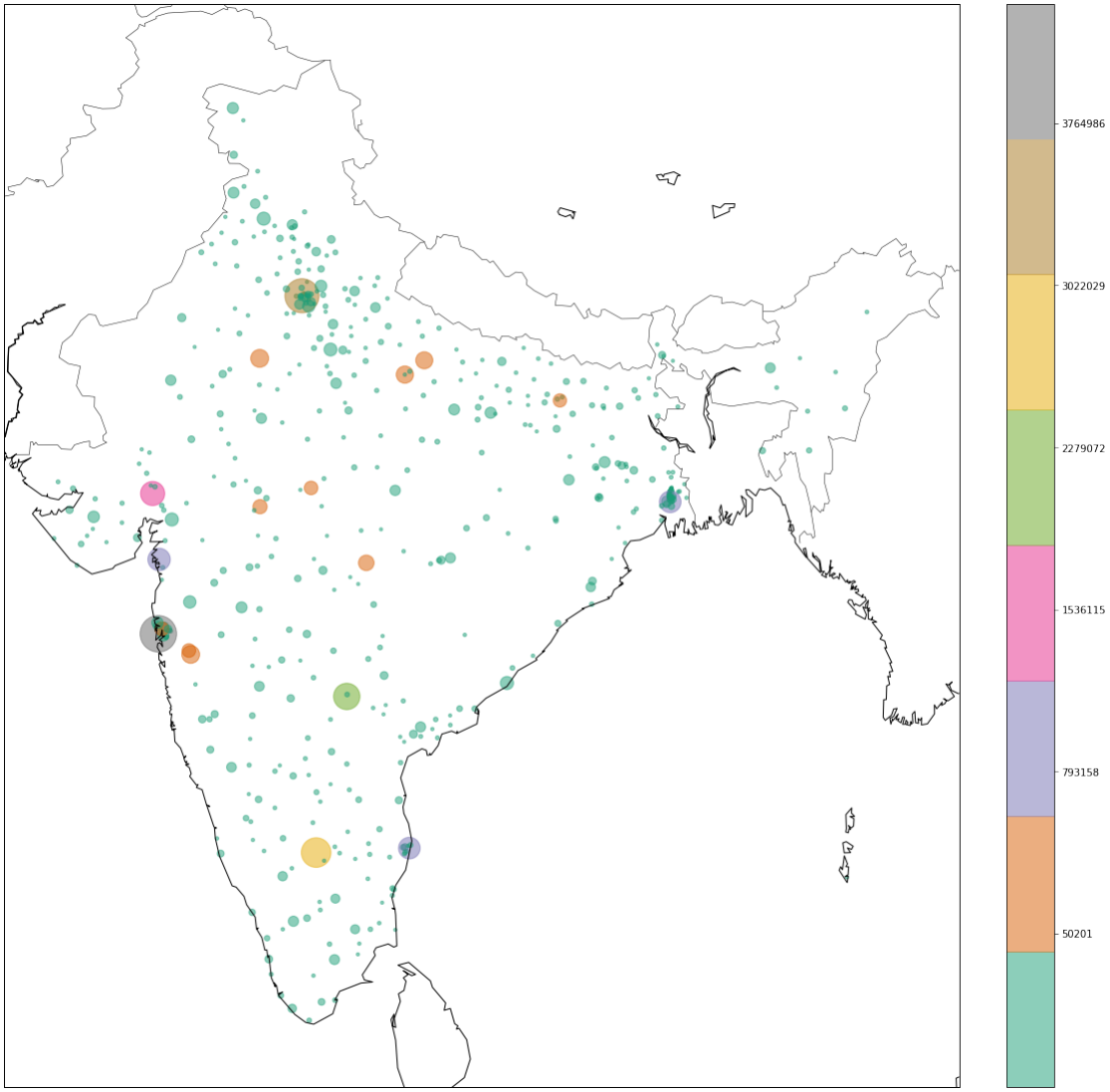
```
In [14]: # A bar chart to show the male population of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['population_male'].sum().sort_values(ascending=
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of cities taken for analysis', fontsize = 20)
         plt.show ()
         # we can see states like Maharashtra and UP have huge male population
```



In [15]:
```
# Plotting the same on the map
population_sizes = cities["population_male"].apply(lambda x: int(x / 5000))
colorbarValue = np.linspace(cities["population_male"].min(), cities["population_male"]
                            num=10)
colorbarValue = colorbarValue.astype(int)

plot_map(population_sizes, colorbarValue)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e644712b0>
```



1.9 These are the top 10 cities with high male population

```
In [16]: # A table to show top 10 cities with most male population
         print("The Top 10 Cities sorted according to the Total Male Population (Descending Or
         top_male_cities = cities.sort_values(by='population_male',ascending=False)
         top10_male_pop_cities=top_male_cities.head(10)
         top10_male_pop_cities
```

```
The Top 10 Cities sorted according to the Total Male Population (Descending Order)
```

Out[16]:

	name_of_city	state_code	state_name	dist_code	\
185	Greater Mumbai	27	MAHARASHTRA	99	
141	Delhi	7	NCT OF DELHI	99	
72	Bengaluru	29	KARNATAKA	18	
184	Greater Hyderabad	28	ANDHRA PRADESH	99	
7	Ahmadabad	24	GUJARAT	7	
449	Surat	24	GUJARAT	25	
274	Kolkata	19	WEST BENGAL	16	
119	Chennai	33	TAMIL NADU	2	
225	Jaipur	8	RAJASTHAN	12	
380	Pune	27	MAHARASHTRA	25	

	population_total	population_male	population_female	\
185	12478447	6736815	5741632	
141	11007835	5871362	5136473	
72	8425970	4401299	4024671	
184	6809970	3500802	3309168	
7	5570585	2935869	2634716	
449	4462002	2538243	1923759	
274	4486679	2362662	2124017	
119	4681087	2357633	2323454	
225	3073350	1619280	1454070	
380	3115431	1602137	1513294	

	0-6_population_total	0-6_population_male	0-6_population_female	...	\
185	1139146	599007	540139	...	
141	1209275	647938	561337	...	
72	862493	444639	417854	...	
184	725816	373794	352022	...	
7	589076	317917	271159	...	
449	531522	293208	238314	...	
274	300052	155475	144577	...	
119	418541	213084	205457	...	
225	378788	204320	174468	...	
380	324572	171152	153420	...	

	child_sex_ratio	effective_literacy_rate_total	\
185	902	90.28	
141	866	87.60	
72	940	89.59	
184	942	82.96	
7	853	89.62	
449	813	89.03	
274	930	87.14	
119	964	90.33	
225	854	84.34	
380	896	91.61	

```
     effective_literacy_rate_male  effective_literacy_rate_female  \
185                         93.32                           86.70
141                         91.44                           83.20
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
449                         92.76                           84.05
274                         89.08                           84.98
119                         93.47                           87.16
225                         90.61                           77.41
380                         95.13                           87.91

                  location  total_graduates  male_graduates  female_graduates  \
185        19.0760,72.8777          1802371          964964            837407
141        28.7041,77.1025          2221137         1210040           1011097
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
449  21.1702401,72.8310607           278795          160566            118229
274    22.572646,88.363895           818476          461615            356861
119  13.0826802,80.2707184           879695          487428            392267
225  26.9124336,75.7872709           533148          319107            214041
380  18.5204303,73.8567437           656508          349022            307486

       latitude   longitude
185     19.0760     72.8777
141     28.7041     77.1025
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
449  21.1702401  72.8310607
274   22.572646   88.363895
119  13.0826802  80.2707184
225  26.9124336  75.7872709
380  18.5204303  73.8567437

[10 rows x 24 columns]
```

In [17]:
```
# Plotting these top 10 male populous cities on India map. Circles are sized accordin
# male population of the respective city

plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()
```

```
        lg=array(top10_male_pop_cities['longitude'])
        lt=array(top10_male_pop_cities['latitude'])
        pt=array(top10_male_pop_cities['population_male'])
        nc=array(top10_male_pop_cities['name_of_city'])

        x, y = map(lg, lt)
        population_sizes_male = top10_male_pop_cities["population_male"].apply(lambda x: int(
        plt.scatter(x, y, s=population_sizes_male, marker="o", c=population_sizes_male, cmap=


        for ncs, xpt, ypt in zip(nc, x, y):
            plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

        plt.title('Top 10 Male Populated Cities in India',fontsize=20)
```

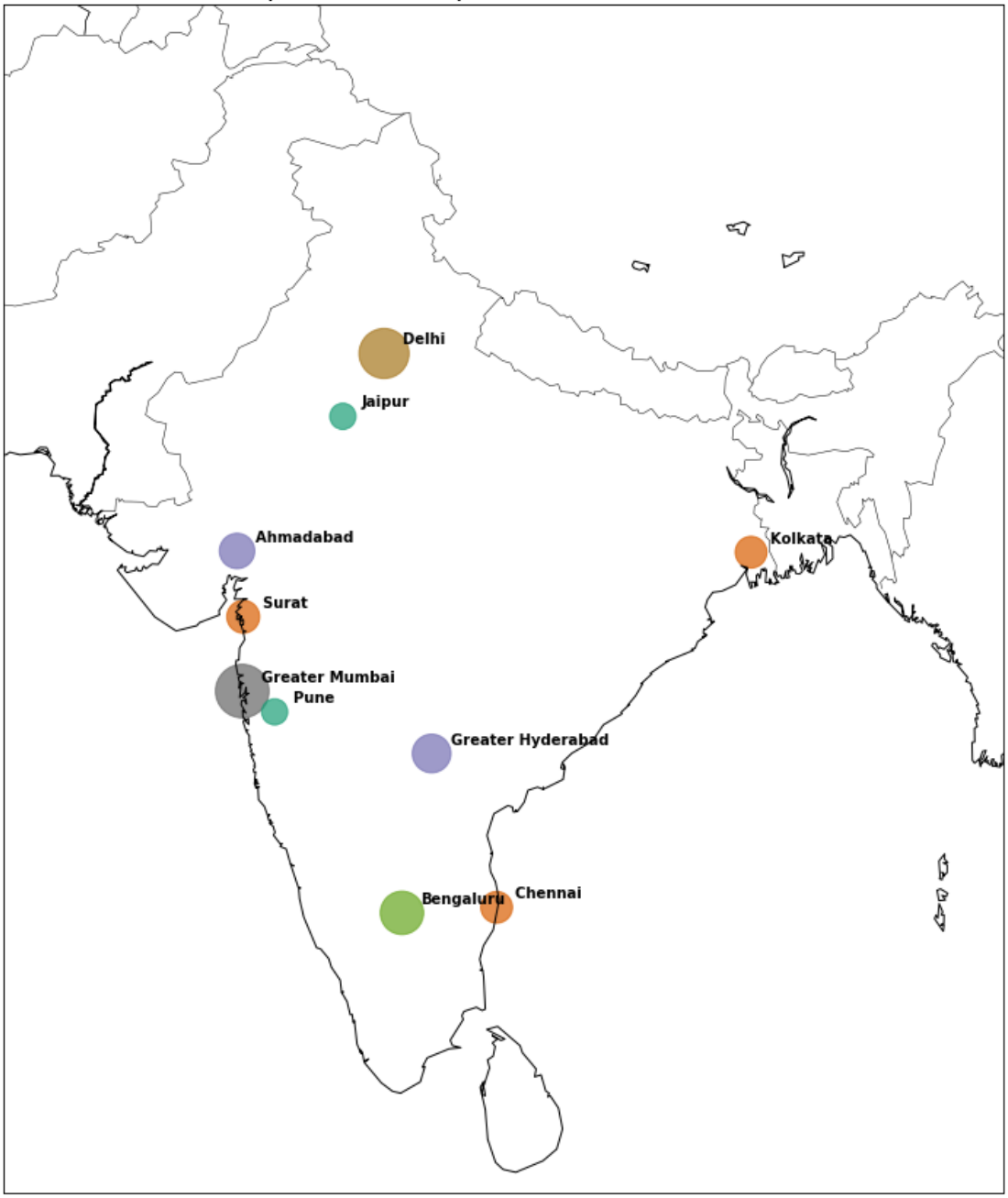
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

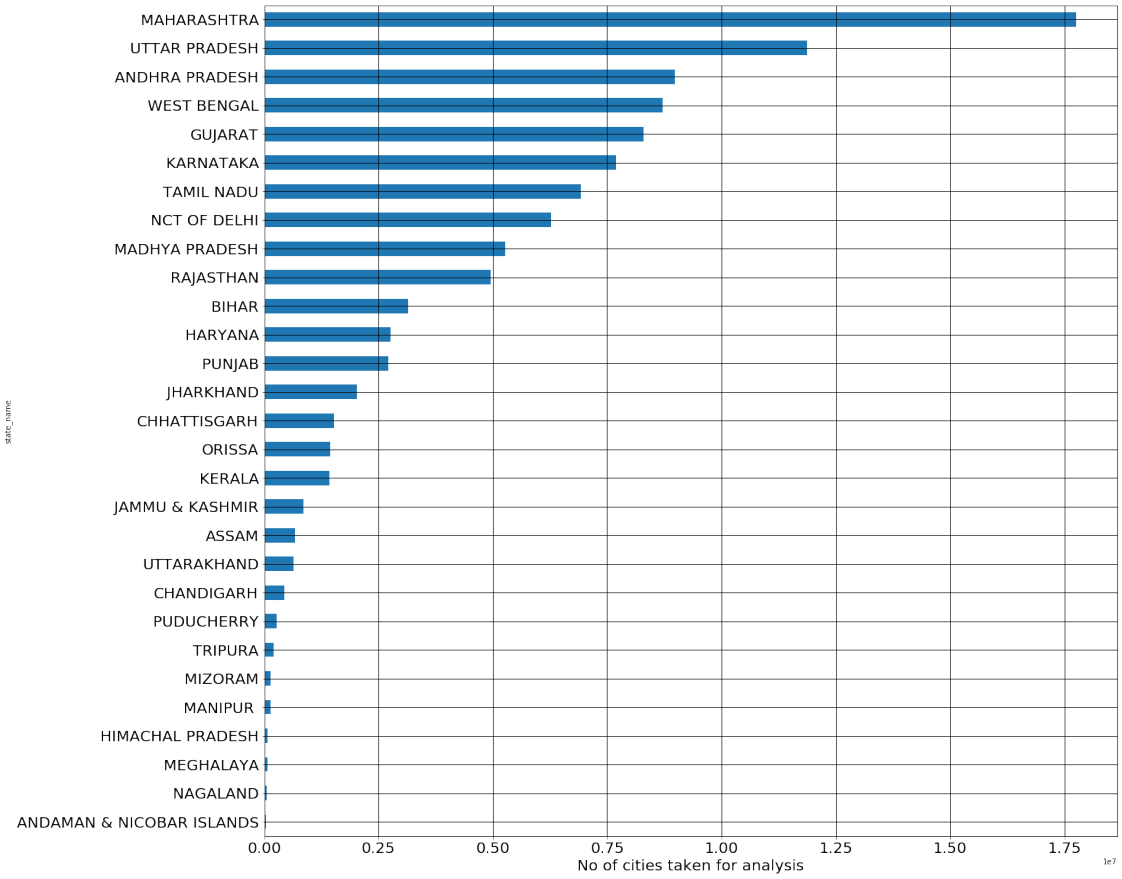
```
Out[17]: Text(0.5,1,'Top 10 Male Populated Cities in India')
```

1.10 Plotting Statewise cities to check which state have most female population

```
In [18]: # A bar chart to show the female population of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['population_female'].sum().sort_values(ascending
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of cities taken for analysis', fontsize = 20)
         plt.show ()
         # we can see again states like Maharashtra and UP have huge female population
```

Top 10 Male Populated Cities in India

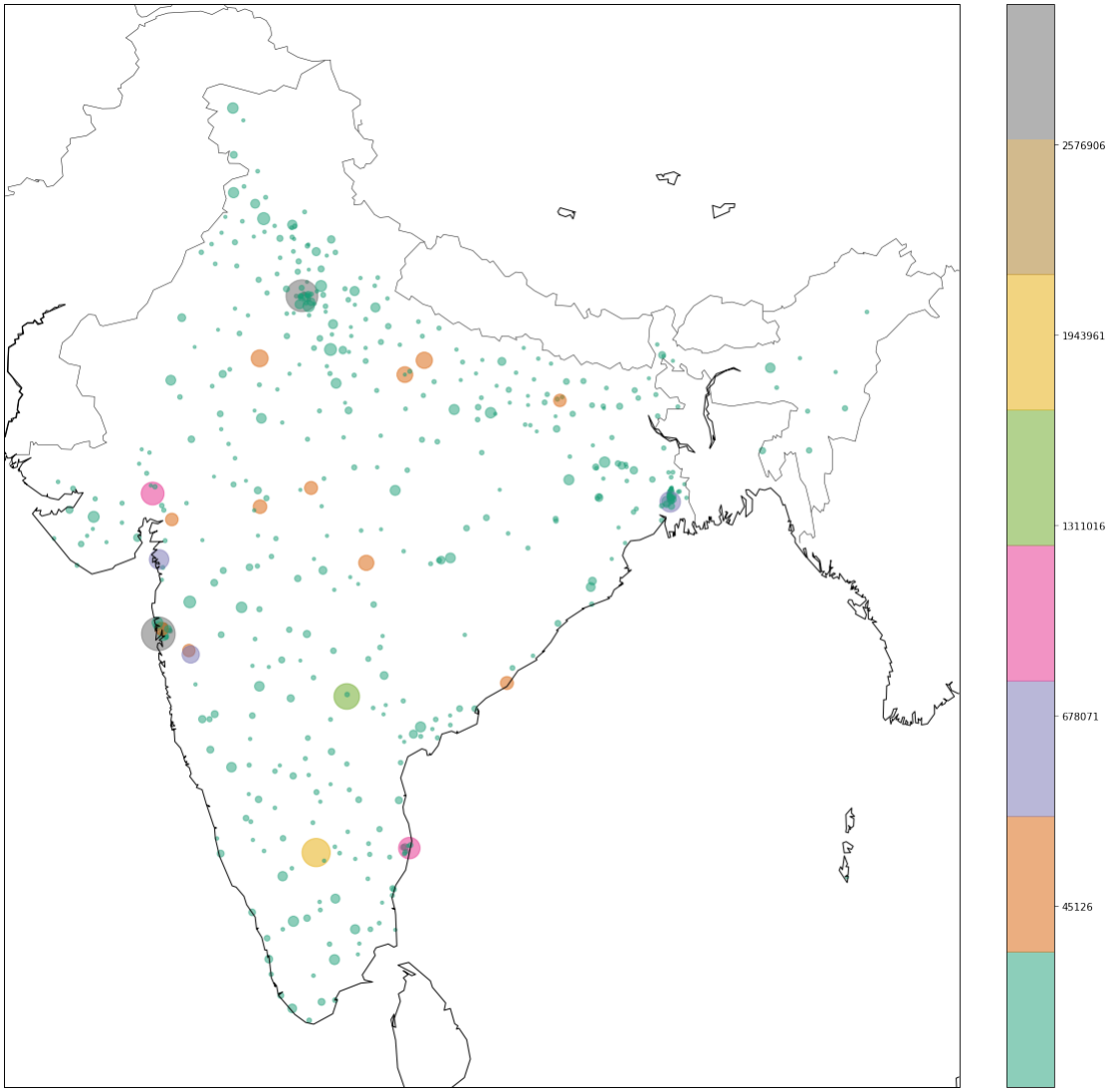




In [19]:
```
# Plotting the same on the map
population_sizes = cities["population_female"].apply(lambda x: int(x / 5000))
colorbarValue = np.linspace(cities["population_female"].min(), cities["population_fema
                            num=10)
colorbarValue = colorbarValue.astype(int)

plot_map(population_sizes, colorbarValue)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e62c1e240>
```



1.11 These are the top 10 cities with high female population

```
In [20]: # A table to show top 10 cities with most female population
         print("The Top 10 Cities sorted according to the Total Female Population (Descending
         top_female_cities = cities.sort_values(by='population_female',ascending=False)
         top10_female_pop_cities=top_female_cities.head(10)
         top10_female_pop_cities
```

```
The Top 10 Cities sorted according to the Total Female Population (Descending Order)
```

Out[20]:

	name_of_city	state_code	state_name	dist_code	\
185	Greater Mumbai	27	MAHARASHTRA	99	
141	Delhi	7	NCT OF DELHI	99	
72	Bengaluru	29	KARNATAKA	18	
184	Greater Hyderabad	28	ANDHRA PRADESH	99	
7	Ahmadabad	24	GUJARAT	7	
119	Chennai	33	TAMIL NADU	2	
274	Kolkata	19	WEST BENGAL	16	
449	Surat	24	GUJARAT	25	
380	Pune	27	MAHARASHTRA	25	
225	Jaipur	8	RAJASTHAN	12	

	population_total	population_male	population_female	\
185	12478447	6736815	5741632	
141	11007835	5871362	5136473	
72	8425970	4401299	4024671	
184	6809970	3500802	3309168	
7	5570585	2935869	2634716	
119	4681087	2357633	2323454	
274	4486679	2362662	2124017	
449	4462002	2538243	1923759	
380	3115431	1602137	1513294	
225	3073350	1619280	1454070	

	0-6_population_total	0-6_population_male	0-6_population_female	...	\
185	1139146	599007	540139	...	
141	1209275	647938	561337	...	
72	862493	444639	417854	...	
184	725816	373794	352022	...	
7	589076	317917	271159	...	
119	418541	213084	205457	...	
274	300052	155475	144577	...	
449	531522	293208	238314	...	
380	324572	171152	153420	...	
225	378788	204320	174468	...	

	child_sex_ratio	effective_literacy_rate_total	\
185	902	90.28	
141	866	87.60	
72	940	89.59	
184	942	82.96	
7	853	89.62	
119	964	90.33	
274	930	87.14	
449	813	89.03	
380	896	91.61	
225	854	84.34	

```
     effective_literacy_rate_male  effective_literacy_rate_female  \
185                         93.32                           86.70
141                         91.44                           83.20
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
119                         93.47                           87.16
274                         89.08                           84.98
449                         92.76                           84.05
380                         95.13                           87.91
225                         90.61                           77.41

                  location  total_graduates  male_graduates  female_graduates  \
185        19.0760,72.8777          1802371          964964            837407
141        28.7041,77.1025          2221137         1210040           1011097
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
119  13.0826802,80.2707184           879695          487428            392267
274    22.572646,88.363895           818476          461615            356861
449  21.1702401,72.8310607           278795          160566            118229
380  18.5204303,73.8567437           656508          349022            307486
225  26.9124336,75.7872709           533148          319107            214041

       latitude   longitude
185     19.0760     72.8777
141     28.7041     77.1025
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
119  13.0826802  80.2707184
274   22.572646   88.363895
449  21.1702401  72.8310607
380  18.5204303  73.8567437
225  26.9124336  75.7872709

[10 rows x 24 columns]
```

In [21]:
```
# Plotting these top 10 female populous cities on India map. Circles are sized accord
# female population of the respective city

plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()
```

```
lg=array(top10_female_pop_cities['longitude'])
lt=array(top10_female_pop_cities['latitude'])
pt=array(top10_female_pop_cities['population_female'])
nc=array(top10_female_pop_cities['name_of_city'])

x, y = map(lg, lt)
population_sizes_female = top10_female_pop_cities["population_female"].apply(lambda x
plt.scatter(x, y, s=population_sizes_female, marker="o", c=population_sizes_female, cm


for ncs, xpt, ypt in zip(nc, x, y):
    plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

plt.title('Top 10 Female Populated Cities in India',fontsize=20)
```

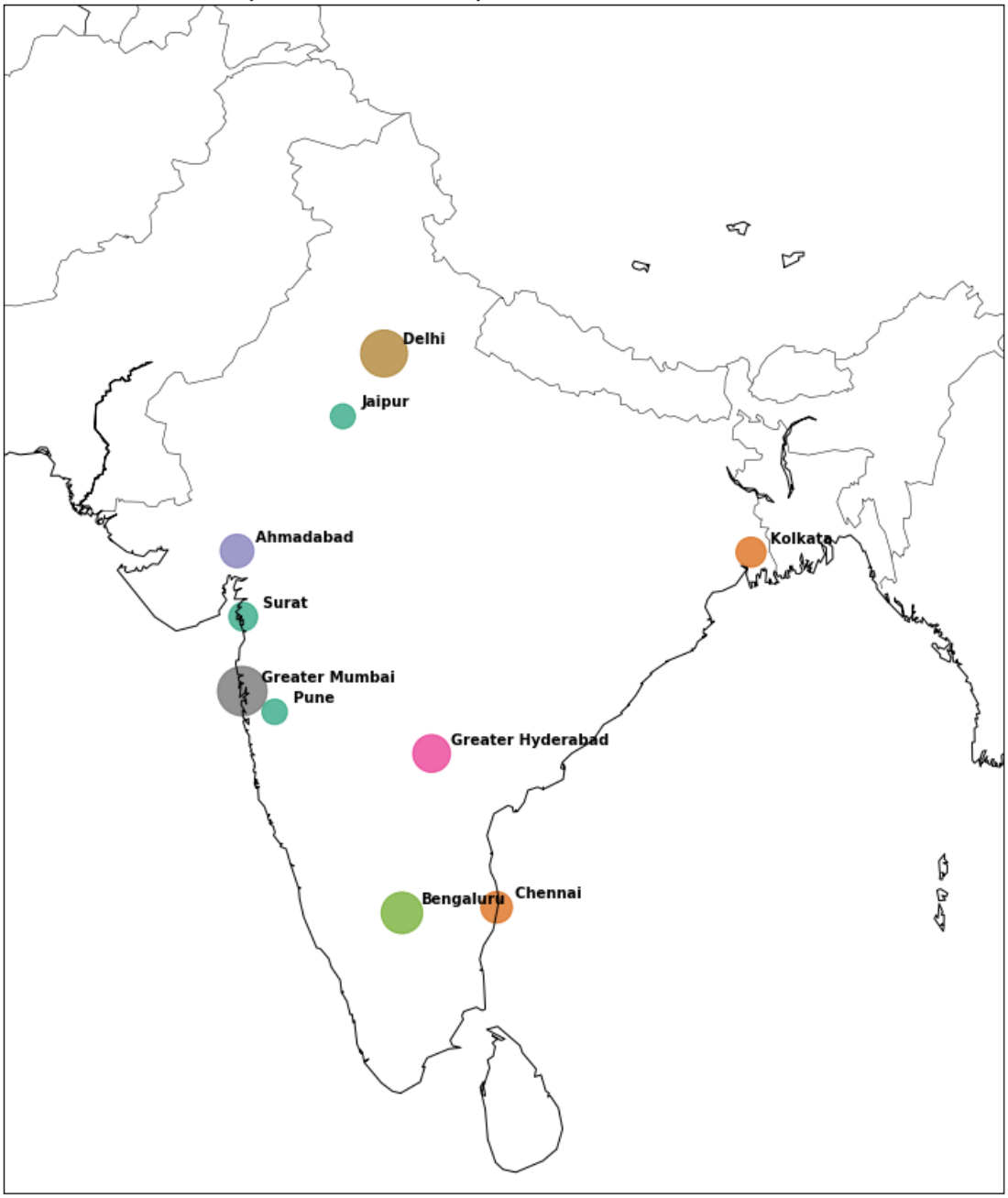
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

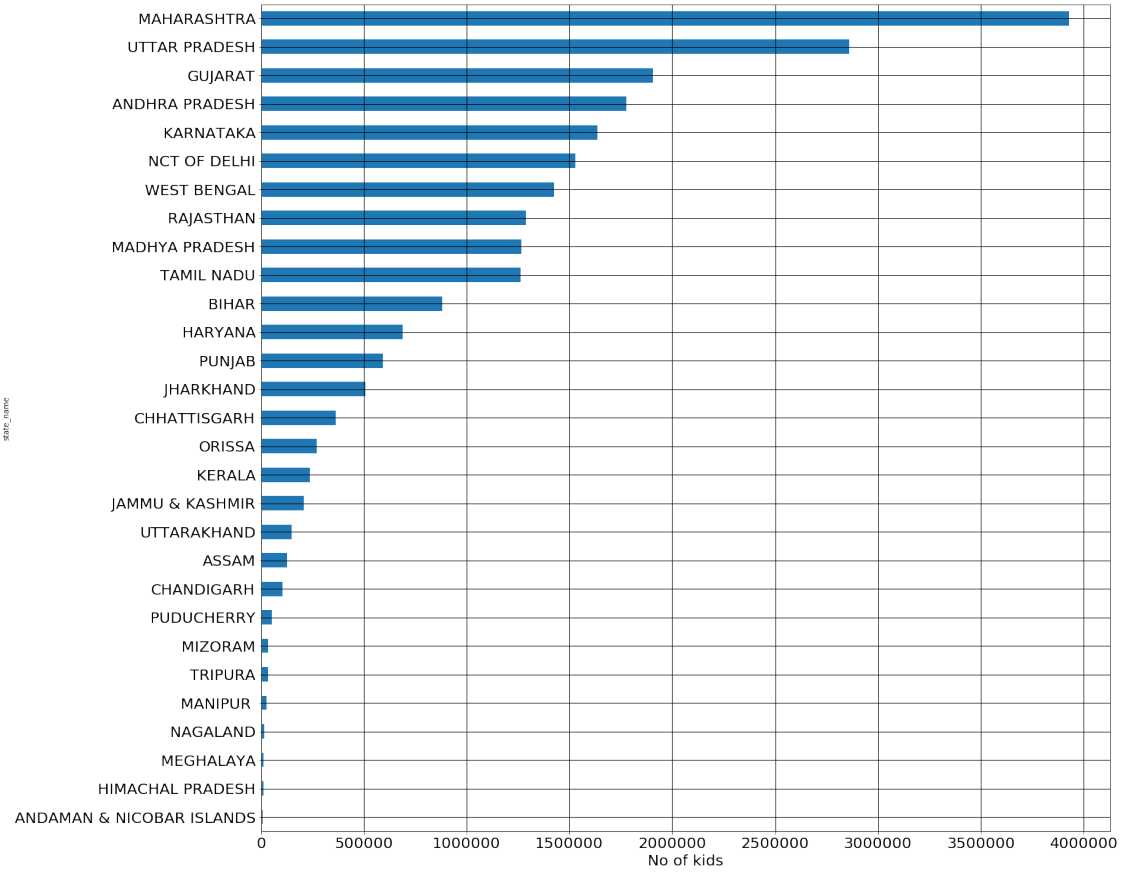
Out[21]: Text(0.5,1,'Top 10 Female Populated Cities in India')

1.12 Plotting Statewise cities to check which state have most Kids (aged between 0 to 6) population

In [22]:
```
# A bar chart to show the kids population of the states
fig = plt.figure(figsize=(20,20))
states = cities.groupby('state_name')['0-6_population_total'].sum().sort_values(ascend
states.plot(kind="barh", fontsize = 20)
plt.grid(b=True, which='both', color='Black',linestyle='-')
plt.xlabel('No of kids', fontsize = 20)
plt.show ()
# we can see again states like Maharashtra and UP have huge kids population living in
```

Top 10 Female Populated Cities in India

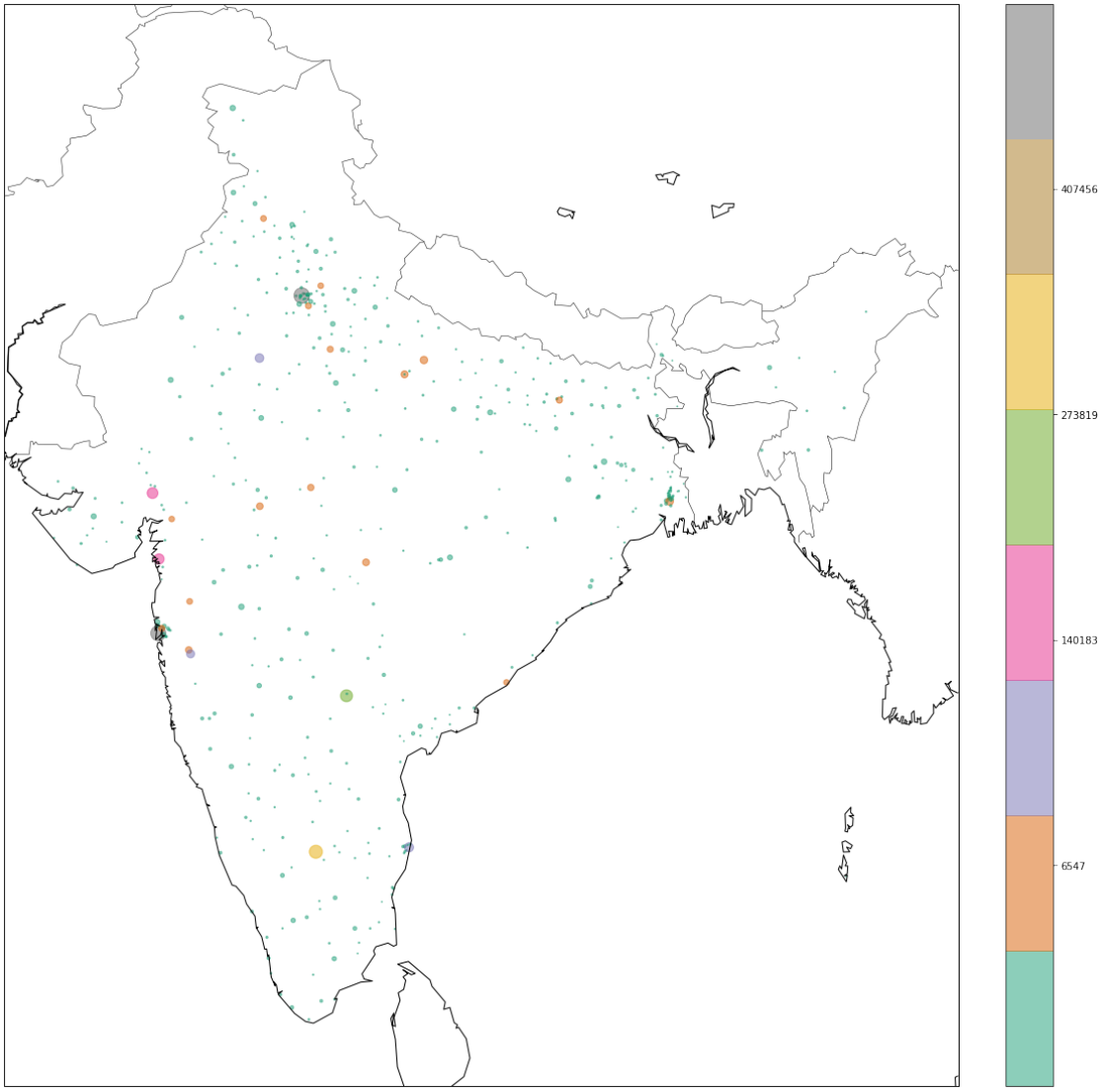




In [23]:
```
# Plotting the same on the map
population_sizes = cities["0-6_population_total"].apply(lambda x: int(x / 5000))
colorbarValue = np.linspace(cities["0-6_population_total"].min(), cities["0-6_populat
                            num=10)
colorbarValue = colorbarValue.astype(int)

plot_map(population_sizes, colorbarValue)
# Kids population is obviously smaller than the overall population and bigger cities
# Mumbai, Banglore, Kolkata, Hyderabad, Chennai have vast number of kids living in ci
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e64475048>
```



1.13 Top 10 cities where most of the kids live

```
In [24]: # Lets find the top ten cities in which large number of kids live
         print("The Top 10 Cities sorted according to the Total Kids Population (Descending Or
         top_kids_cities = cities.sort_values(by='0-6_population_total',ascending=False)
         top10_kids_pop_cities=top_kids_cities.head(10)
         top10_kids_pop_cities
```

```
The Top 10 Cities sorted according to the Total Kids Population (Descending Order)
```

Out[24]:

	name_of_city	state_code	state_name	dist_code	\
141	Delhi	7	NCT OF DELHI	99	
185	Greater Mumbai	27	MAHARASHTRA	99	
72	Bengaluru	29	KARNATAKA	18	
184	Greater Hyderabad	28	ANDHRA PRADESH	99	
7	Ahmadabad	24	GUJARAT	7	
449	Surat	24	GUJARAT	25	
119	Chennai	33	TAMIL NADU	2	
225	Jaipur	8	RAJASTHAN	12	
380	Pune	27	MAHARASHTRA	25	
274	Kolkata	19	WEST BENGAL	16	

	population_total	population_male	population_female	\
141	11007835	5871362	5136473	
185	12478447	6736815	5741632	
72	8425970	4401299	4024671	
184	6809970	3500802	3309168	
7	5570585	2935869	2634716	
449	4462002	2538243	1923759	
119	4681087	2357633	2323454	
225	3073350	1619280	1454070	
380	3115431	1602137	1513294	
274	4486679	2362662	2124017	

	0-6_population_total	0-6_population_male	0-6_population_female	...	\
141	1209275	647938	561337	...	
185	1139146	599007	540139	...	
72	862493	444639	417854	...	
184	725816	373794	352022	...	
7	589076	317917	271159	...	
449	531522	293208	238314	...	
119	418541	213084	205457	...	
225	378788	204320	174468	...	
380	324572	171152	153420	...	
274	300052	155475	144577	...	

	child_sex_ratio	effective_literacy_rate_total	\
141	866	87.60	
185	902	90.28	
72	940	89.59	
184	942	82.96	
7	853	89.62	
449	813	89.03	
119	964	90.33	
225	854	84.34	
380	896	91.61	
274	930	87.14	

```
     effective_literacy_rate_male  effective_literacy_rate_female  \
141                         91.44                           83.20
185                         93.32                           86.70
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
449                         92.76                           84.05
119                         93.47                           87.16
225                         90.61                           77.41
380                         95.13                           87.91
274                         89.08                           84.98

                  location  total_graduates  male_graduates  female_graduates  \
141        28.7041,77.1025          2221137         1210040           1011097
185        19.0760,72.8777          1802371          964964            837407
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
449  21.1702401,72.8310607           278795          160566            118229
119  13.0826802,80.2707184           879695          487428            392267
225  26.9124336,75.7872709           533148          319107            214041
380  18.5204303,73.8567437           656508          349022            307486
274    22.572646,88.363895           818476          461615            356861

       latitude   longitude
141     28.7041     77.1025
185     19.0760     72.8777
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
449  21.1702401  72.8310607
119  13.0826802  80.2707184
225  26.9124336  75.7872709
380  18.5204303  73.8567437
274   22.572646   88.363895

[10 rows x 24 columns]
```

In [25]:
```
# Lets find the top ten cities in which large number of kids live

plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()
```

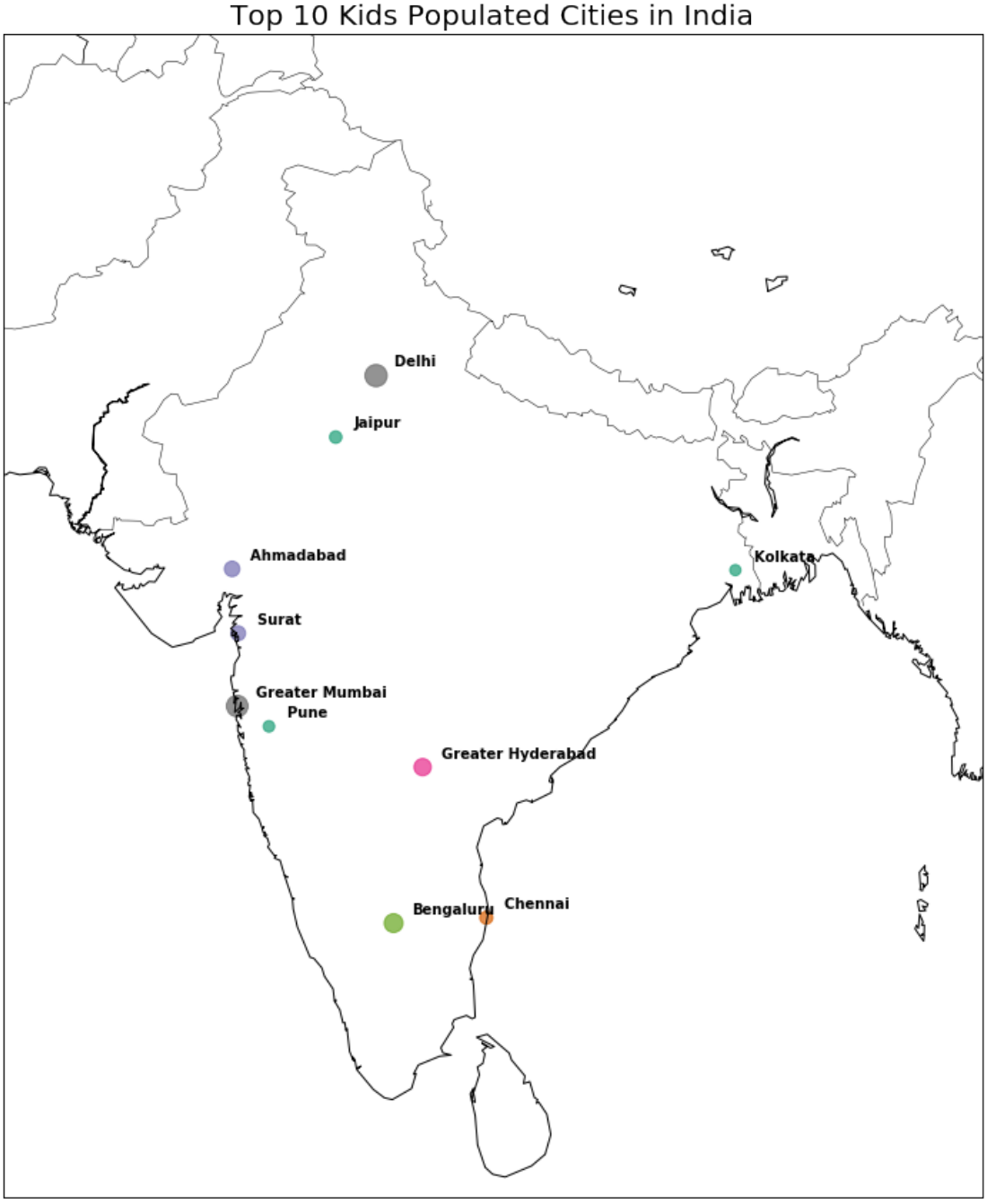
```
        lg=array(top10_kids_pop_cities['longitude'])
        lt=array(top10_kids_pop_cities['latitude'])
        pt=array(top10_kids_pop_cities['0-6_population_total'])
        nc=array(top10_kids_pop_cities['name_of_city'])

        x, y = map(lg, lt)
        population_sizes_kids = top10_kids_pop_cities["0-6_population_total"].apply(lambda x:
        plt.scatter(x, y, s=population_sizes_kids, marker="o", c=population_sizes_kids, cmap=


        for ncs, xpt, ypt in zip(nc, x, y):
            plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

        plt.title('Top 10 Kids Populated Cities in India',fontsize=20)
```

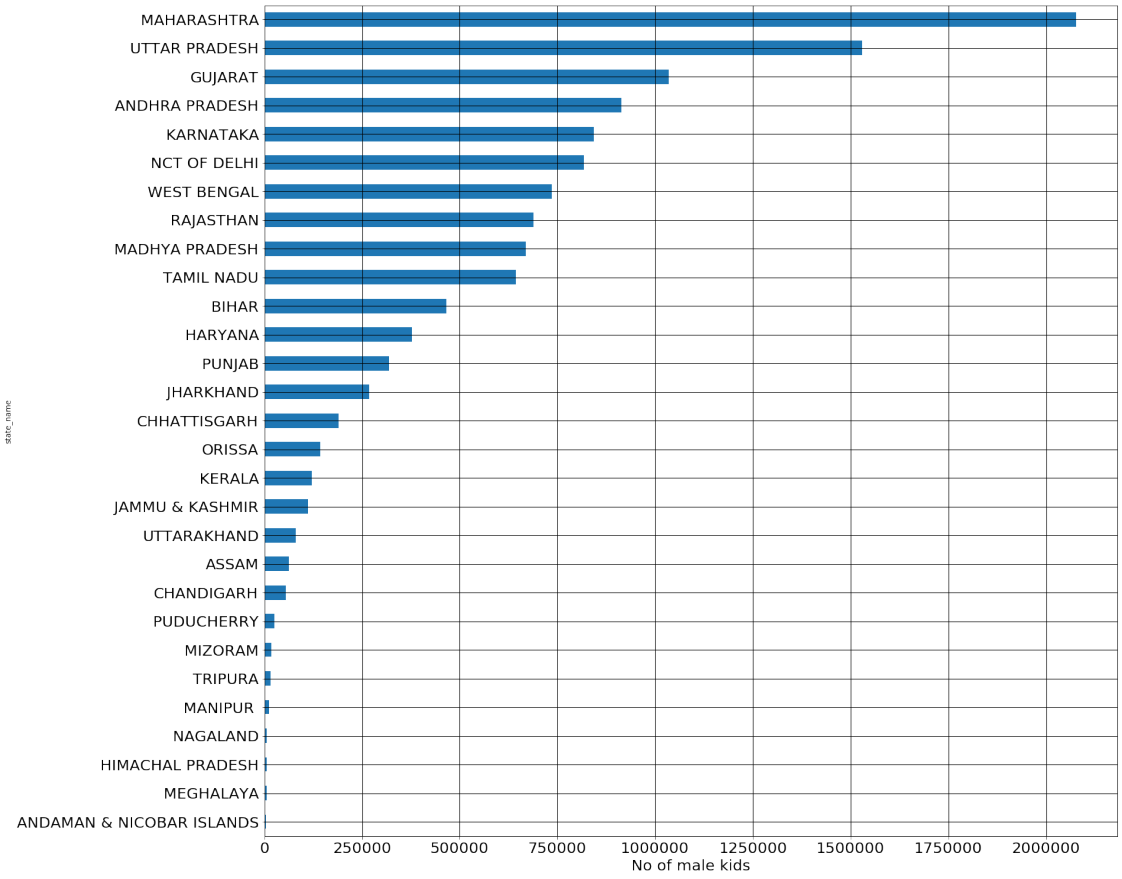
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

```
Out[25]: Text(0.5,1,'Top 10 Kids Populated Cities in India')
```

1.14 Plotting Statewise cities to check which state have most Male Kids (aged between 0 to 6) population

```
In [26]: # A bar chart to show the male kids population of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['0-6_population_male'].sum().sort_values(ascendi
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of male kids', fontsize = 20)
         plt.show ()
         # we can see again states like Maharashtra and UP have huge male kids population livi
```

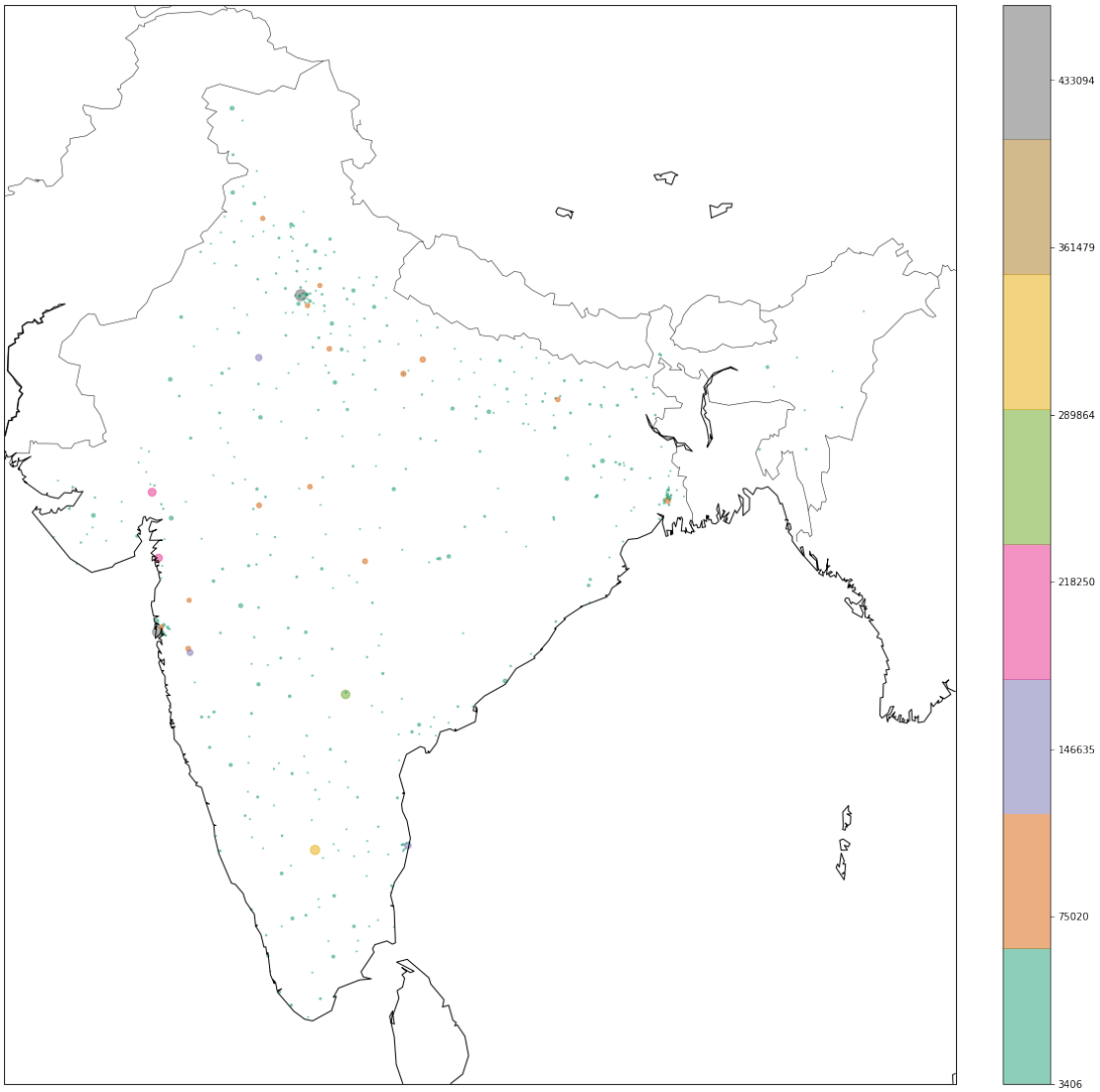

Top 10 Kids Populated Cities in India
Delhi
Jaipur
Ahmadabad
Kolkata
Surat
Greater Mumbai
Pune
Greater Hyderabad
Bengaluru
Chennai



In [27]:
```
# Plotting the same on the map
population_sizes = cities["0-6_population_male"].apply(lambda x: int(x / 5000))
colorbarValue = np.linspace(cities["0-6_population_male"].min(), cities["0-6_populatio
                            num=10)
colorbarValue = colorbarValue.astype(int)

plot_map(population_sizes, colorbarValue)
# Kids population is obviously smaller than the overall population and bigger cities
# Mumbai, Banglore, Kolkata, Hyderabad, Chennai have vast number of kids living in ci
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e63ce1a20>
```



1.15 Top 10 cities where most of the male kids live

```
In [28]: # Lets find the top ten cities in which large number of male kids live
         print("The Top 10 Cities sorted according to the Total Male Kids Population (Descendi
         top10_male_kids_cities = cities.sort_values(by='0-6_population_male',ascending=False)
         top10_male_kids_pop_cities=top10_male_kids_cities.head(10)
         top10_male_kids_pop_cities
```

```
The Top 10 Cities sorted according to the Total Male Kids Population (Descending Order)
```

Out[28]:

	name_of_city	state_code	state_name	dist_code	\
141	Delhi	7	NCT OF DELHI	99	
185	Greater Mumbai	27	MAHARASHTRA	99	
72	Bengaluru	29	KARNATAKA	18	
184	Greater Hyderabad	28	ANDHRA PRADESH	99	
7	Ahmadabad	24	GUJARAT	7	
449	Surat	24	GUJARAT	25	
119	Chennai	33	TAMIL NADU	2	
225	Jaipur	8	RAJASTHAN	12	
380	Pune	27	MAHARASHTRA	25	
274	Kolkata	19	WEST BENGAL	16	

	population_total	population_male	population_female	\
141	11007835	5871362	5136473	
185	12478447	6736815	5741632	
72	8425970	4401299	4024671	
184	6809970	3500802	3309168	
7	5570585	2935869	2634716	
449	4462002	2538243	1923759	
119	4681087	2357633	2323454	
225	3073350	1619280	1454070	
380	3115431	1602137	1513294	
274	4486679	2362662	2124017	

	0-6_population_total	0-6_population_male	0-6_population_female	...	\
141	1209275	647938	561337	...	
185	1139146	599007	540139	...	
72	862493	444639	417854	...	
184	725816	373794	352022	...	
7	589076	317917	271159	...	
449	531522	293208	238314	...	
119	418541	213084	205457	...	
225	378788	204320	174468	...	
380	324572	171152	153420	...	
274	300052	155475	144577	...	

	child_sex_ratio	effective_literacy_rate_total	\
141	866	87.60	
185	902	90.28	
72	940	89.59	
184	942	82.96	
7	853	89.62	
449	813	89.03	
119	964	90.33	
225	854	84.34	
380	896	91.61	
274	930	87.14	

```
     effective_literacy_rate_male  effective_literacy_rate_female  \
141                         91.44                           83.20
185                         93.32                           86.70
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
449                         92.76                           84.05
119                         93.47                           87.16
225                         90.61                           77.41
380                         95.13                           87.91
274                         89.08                           84.98

                  location  total_graduates  male_graduates  female_graduates  \
141        28.7041,77.1025          2221137         1210040           1011097
185        19.0760,72.8777          1802371          964964            837407
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
449  21.1702401,72.8310607           278795          160566            118229
119  13.0826802,80.2707184           879695          487428            392267
225  26.9124336,75.7872709           533148          319107            214041
380  18.5204303,73.8567437           656508          349022            307486
274    22.572646,88.363895           818476          461615            356861

       latitude   longitude
141     28.7041     77.1025
185     19.0760     72.8777
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
449  21.1702401  72.8310607
119  13.0826802  80.2707184
225  26.9124336  75.7872709
380  18.5204303  73.8567437
274   22.572646   88.363895

[10 rows x 24 columns]
```

In [29]:
```
# Lets find the top ten cities in which large number of male kids live

plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()
```

```
        lg=array(top10_male_kids_pop_cities['longitude'])
        lt=array(top10_male_kids_pop_cities['latitude'])
        pt=array(top10_male_kids_pop_cities['0-6_population_male'])
        nc=array(top10_male_kids_pop_cities['name_of_city'])

        x, y = map(lg, lt)
        population_sizes_male_kids = top10_male_kids_pop_cities["0-6_population_male"].apply(
        plt.scatter(x, y, s=population_sizes_male_kids, marker="o", c=population_sizes_male_ki


        for ncs, xpt, ypt in zip(nc, x, y):
            plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

        plt.title('Top 10 Male Kids Populated Cities in India',fontsize=20)
```

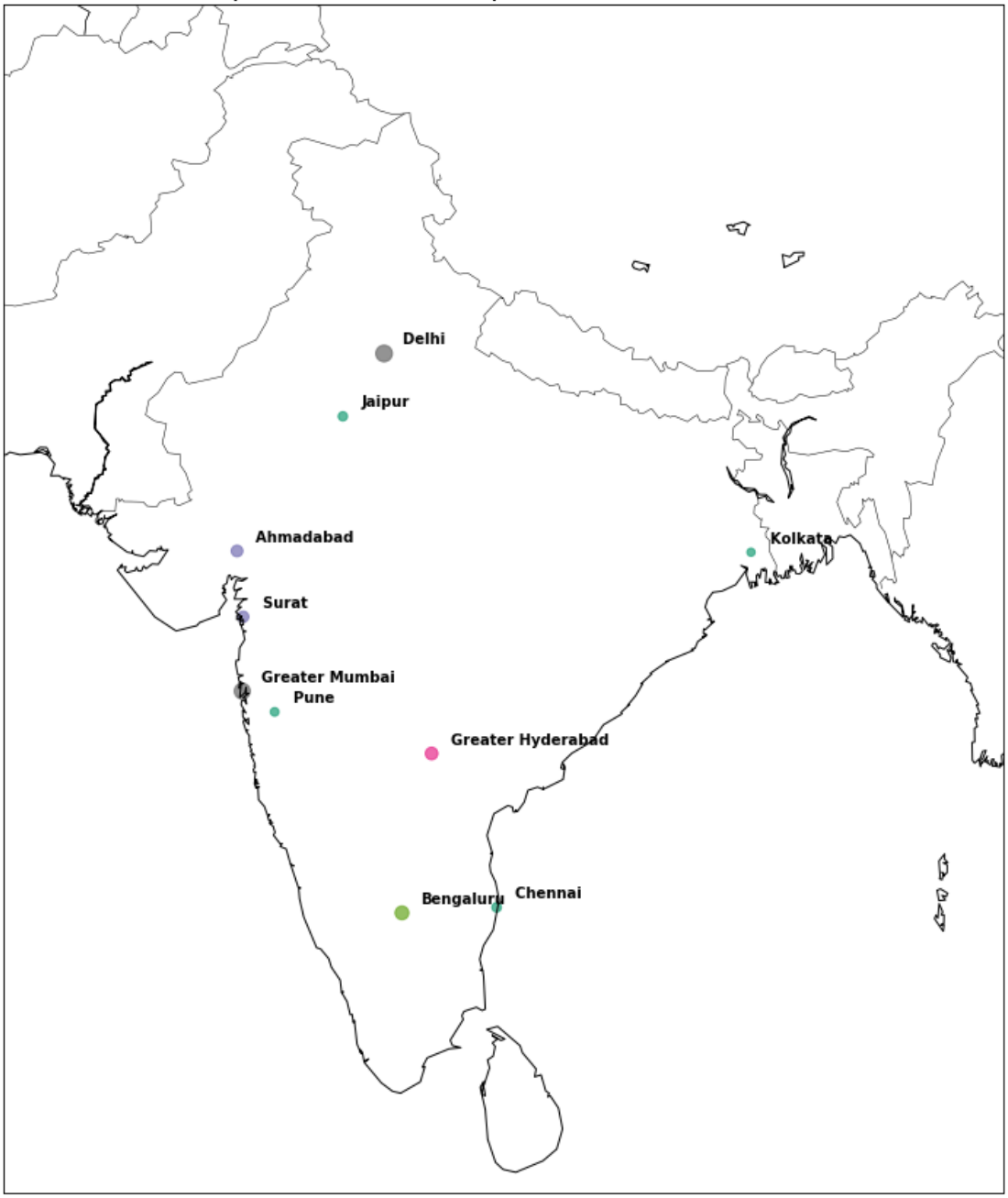
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

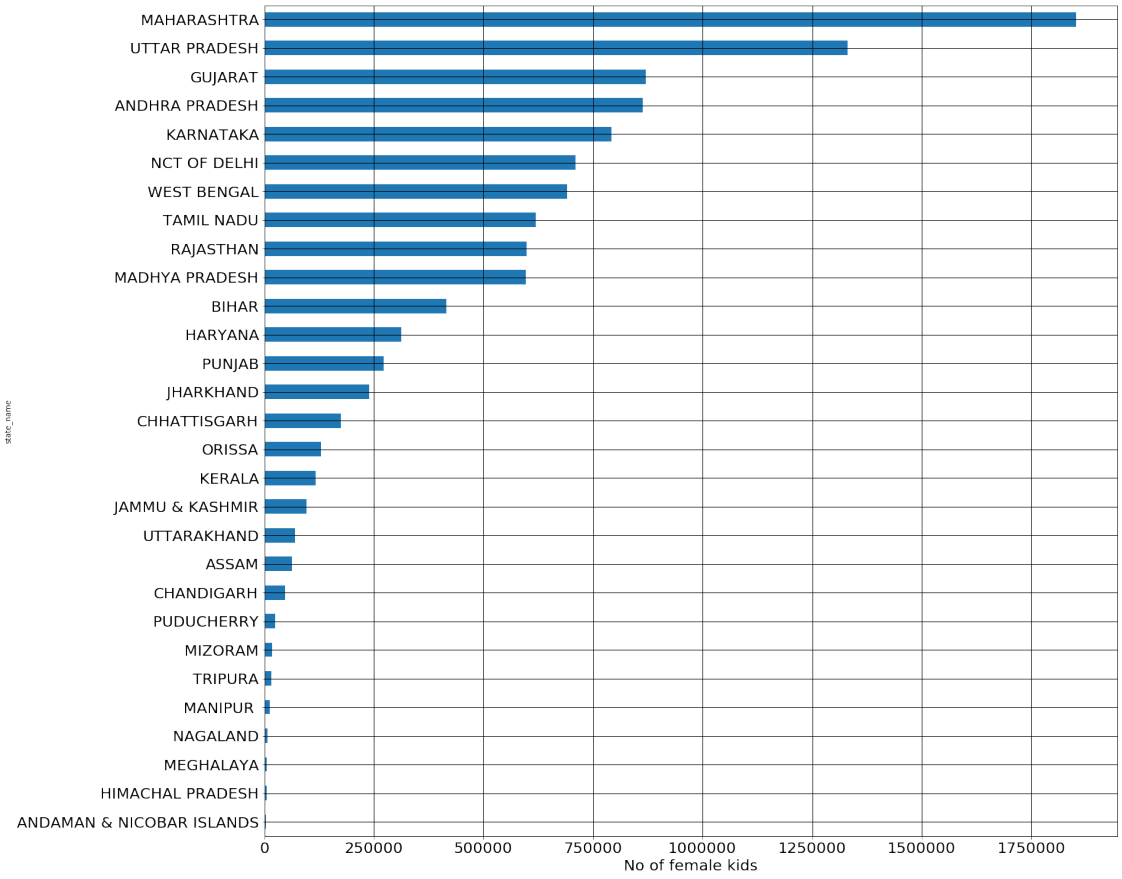
```
Out[29]: Text(0.5,1,'Top 10 Male Kids Populated Cities in India')
```

1.16 Plotting Statewise cities to check which state have most Female Kids (aged between 0 to 6) population

```
In [30]: # A bar chart to show the female kids population of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['0-6_population_female'].sum().sort_values(ascen
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of female kids', fontsize = 20)
         plt.show ()
         # we can see again states like Maharashtra and UP have huge male kids population livi
```

Top 10 Male Kids Populated Cities in India

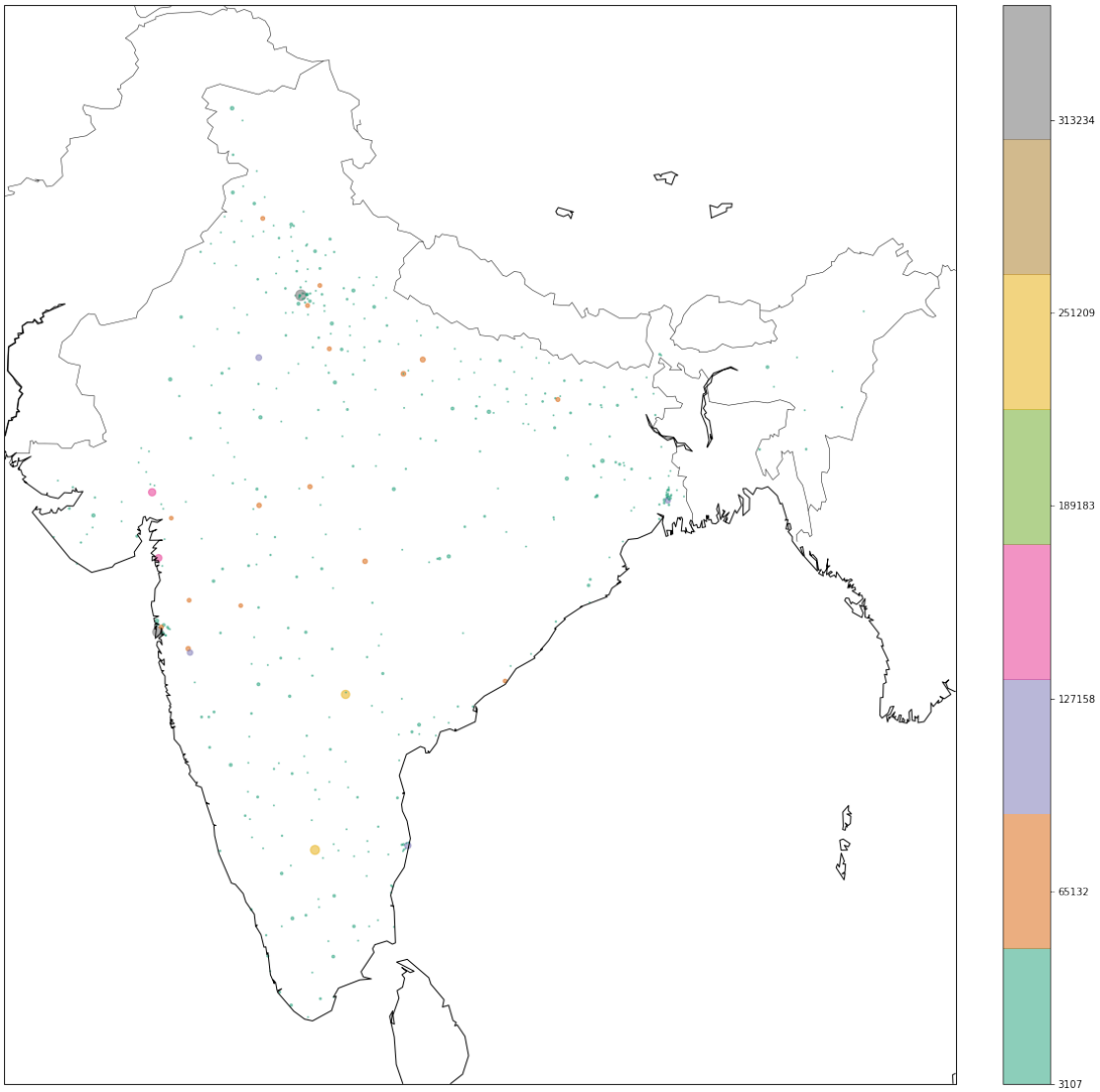




In [31]:
```
# Plotting the same on the map
population_sizes = cities["0-6_population_female"].apply(lambda x: int(x / 5000))
colorbarValue = np.linspace(cities["0-6_population_female"].min(), cities["0-6_popula
                            num=10)
colorbarValue = colorbarValue.astype(int)

plot_map(population_sizes, colorbarValue)
# Kids population is obviously smaller than the overall population and bigger cities
# Mumbai, Banglore, Kolkata, Hyderabad, Chennai have vast number of kids living in ci
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e62d7fa20>
```



1.17 Top 10 cities where most of the female children live

```
In [32]: # Lets find the top ten cities in which large number of female kids live
         print("The Top 10 Cities sorted according to the Total Female Kids Population (Descend
         top10_female_kids_cities = cities.sort_values(by='0-6_population_female',ascending=Fa
         top10_female_kids_pop_cities=top10_female_kids_cities.head(10)
         top10_female_kids_pop_cities
```

```
The Top 10 Cities sorted according to the Total Female Kids Population (Descending Order)
```

Out[32]:

	name_of_city	state_code	state_name	dist_code	\
141	Delhi	7	NCT OF DELHI	99	
185	Greater Mumbai	27	MAHARASHTRA	99	
72	Bengaluru	29	KARNATAKA	18	
184	Greater Hyderabad	28	ANDHRA PRADESH	99	
7	Ahmadabad	24	GUJARAT	7	
449	Surat	24	GUJARAT	25	
119	Chennai	33	TAMIL NADU	2	
225	Jaipur	8	RAJASTHAN	12	
380	Pune	27	MAHARASHTRA	25	
274	Kolkata	19	WEST BENGAL	16	

	population_total	population_male	population_female	\
141	11007835	5871362	5136473	
185	12478447	6736815	5741632	
72	8425970	4401299	4024671	
184	6809970	3500802	3309168	
7	5570585	2935869	2634716	
449	4462002	2538243	1923759	
119	4681087	2357633	2323454	
225	3073350	1619280	1454070	
380	3115431	1602137	1513294	
274	4486679	2362662	2124017	

	0-6_population_total	0-6_population_male	0-6_population_female	...	\
141	1209275	647938	561337	...	
185	1139146	599007	540139	...	
72	862493	444639	417854	...	
184	725816	373794	352022	...	
7	589076	317917	271159	...	
449	531522	293208	238314	...	
119	418541	213084	205457	...	
225	378788	204320	174468	...	
380	324572	171152	153420	...	
274	300052	155475	144577	...	

	child_sex_ratio	effective_literacy_rate_total	\
141	866	87.60	
185	902	90.28	
72	940	89.59	
184	942	82.96	
7	853	89.62	
449	813	89.03	
119	964	90.33	
225	854	84.34	
380	896	91.61	
274	930	87.14	

```
     effective_literacy_rate_male  effective_literacy_rate_female  \
141                         91.44                           83.20
185                         93.32                           86.70
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
449                         92.76                           84.05
119                         93.47                           87.16
225                         90.61                           77.41
380                         95.13                           87.91
274                         89.08                           84.98

                  location  total_graduates  male_graduates  female_graduates  \
141        28.7041,77.1025          2221137         1210040           1011097
185        19.0760,72.8777          1802371          964964            837407
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
449  21.1702401,72.8310607           278795          160566            118229
119  13.0826802,80.2707184           879695          487428            392267
225  26.9124336,75.7872709           533148          319107            214041
380  18.5204303,73.8567437           656508          349022            307486
274    22.572646,88.363895           818476          461615            356861

       latitude   longitude
141     28.7041     77.1025
185     19.0760     72.8777
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
449  21.1702401  72.8310607
119  13.0826802  80.2707184
225  26.9124336  75.7872709
380  18.5204303  73.8567437
274   22.572646   88.363895

[10 rows x 24 columns]
```

In [33]:
```
# Lets find the top ten cities in which large number of female kids live

plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()
```

```
        lg=array(top10_female_kids_pop_cities['longitude'])
        lt=array(top10_female_kids_pop_cities['latitude'])
        pt=array(top10_female_kids_pop_cities['0-6_population_female'])
        nc=array(top10_female_kids_pop_cities['name_of_city'])

        x, y = map(lg, lt)
        population_sizes_female_kids = top10_female_kids_pop_cities["0-6_population_female"].a
        plt.scatter(x, y, s=population_sizes_female_kids, marker="o", c=population_sizes_femal


        for ncs, xpt, ypt in zip(nc, x, y):
            plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

        plt.title('Top 10 Female Kids Populated Cities in India',fontsize=20)
```

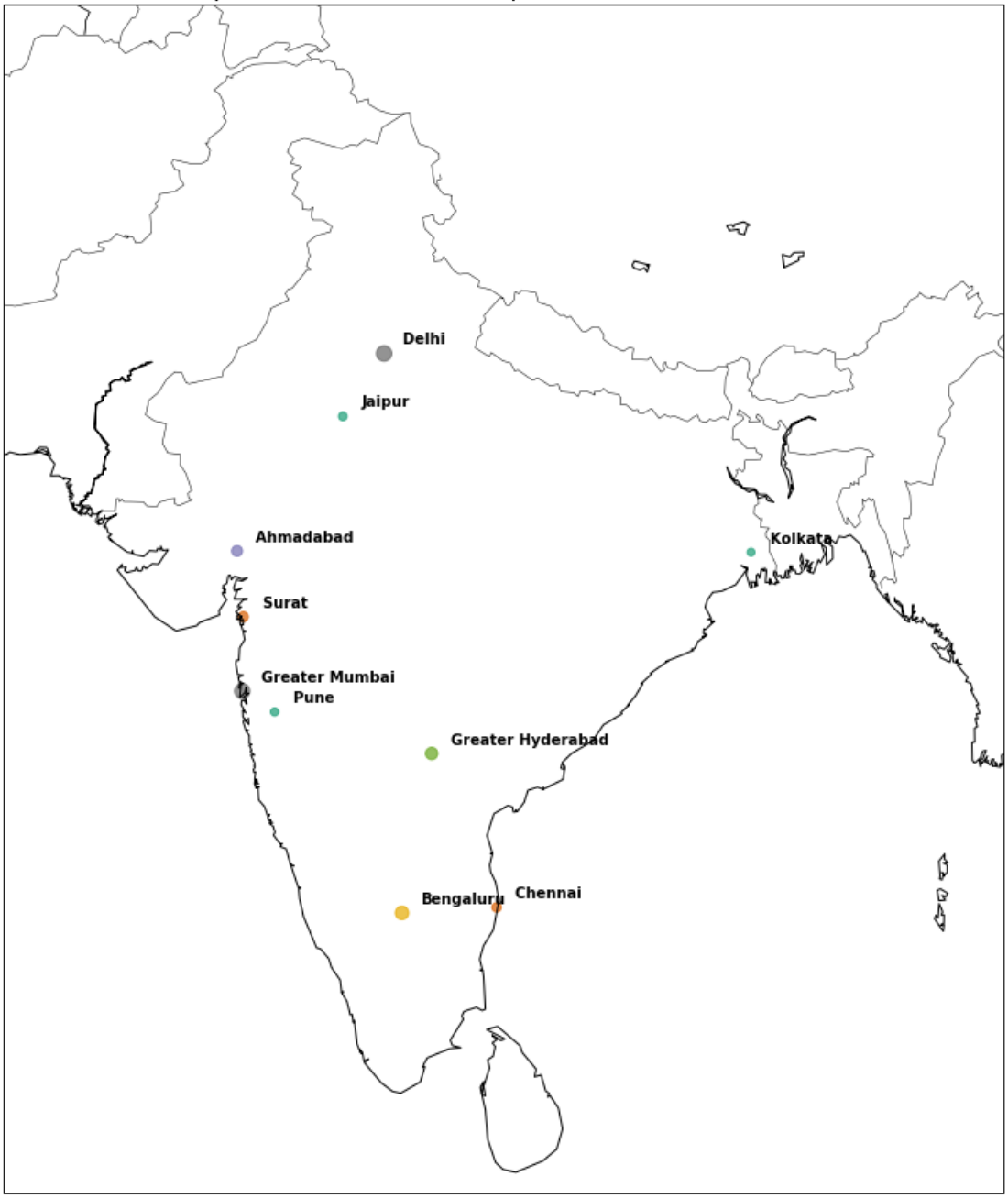
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

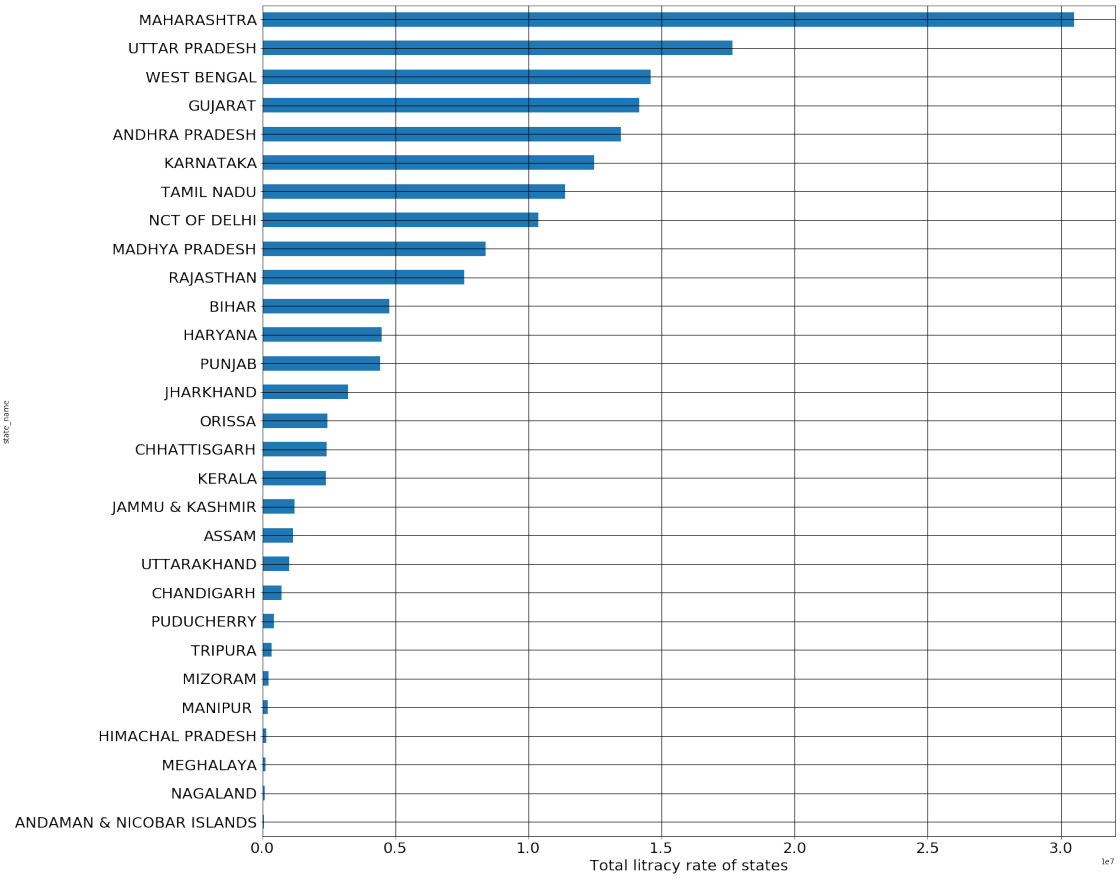
```
Out[33]: Text(0.5,1,'Top 10 Female Kids Populated Cities in India')
```

1.18 Analysing Litracy rate of the states

```
In [34]: # A bar chart to show the total literates of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['literates_total'].sum().sort_values(ascending='
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('Total litracy rate of states', fontsize = 20)
         plt.show ()
         # we can see again states like Maharashtra and UP have huge litrate population living
```

Top 10 Female Kids Populated Cities in India

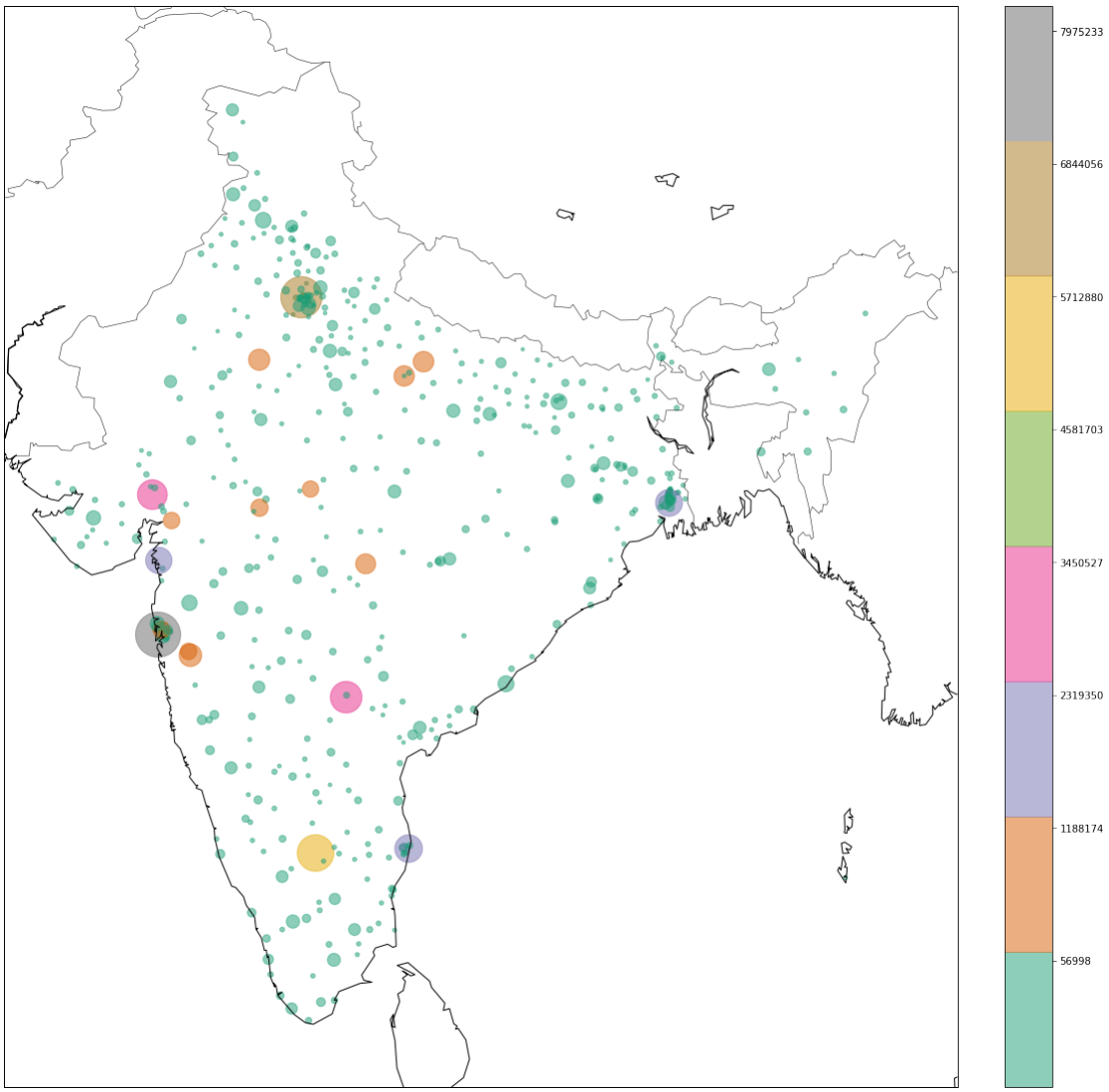




In [35]:
```
# Plotting the same on the map
population_sizes = cities["literates_total"].apply(lambda x: int(x / 5000))
colorbarValue = np.linspace(cities["literates_total"].min(), cities["literates_total"]
                            num=10)
colorbarValue = colorbarValue.astype(int)

plot_map(population_sizes, colorbarValue)
# Major metro cities again shows higher litracy rates
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e643e8198>
```



1.19 Top 10 cities where most of the literates live

```
In [36]: # Lets find the top ten cities in which large number of literates live
         print("The Top 10 Cities sorted according to the Total litrate Population (Descending
         top10_literate_cities = cities.sort_values(by='literates_total',ascending=False)
         top10_literate_cities=top10_literate_cities.head(10)
         top10_literate_cities
```

```
The Top 10 Cities sorted according to the Total litrate Population (Descending Order)
```

Out[36]:

	name_of_city	state_code	state_name	dist_code	\
185	Greater Mumbai	27	MAHARASHTRA	99	
141	Delhi	7	NCT OF DELHI	99	
72	Bengaluru	29	KARNATAKA	18	
184	Greater Hyderabad	28	ANDHRA PRADESH	99	
7	Ahmadabad	24	GUJARAT	7	
119	Chennai	33	TAMIL NADU	2	
274	Kolkata	19	WEST BENGAL	16	
449	Surat	24	GUJARAT	25	
380	Pune	27	MAHARASHTRA	25	
225	Jaipur	8	RAJASTHAN	12	

	population_total	population_male	population_female	\
185	12478447	6736815	5741632	
141	11007835	5871362	5136473	
72	8425970	4401299	4024671	
184	6809970	3500802	3309168	
7	5570585	2935869	2634716	
119	4681087	2357633	2323454	
274	4486679	2362662	2124017	
449	4462002	2538243	1923759	
380	3115431	1602137	1513294	
225	3073350	1619280	1454070	

	0-6_population_total	0-6_population_male	0-6_population_female	...	\
185	1139146	599007	540139	...	
141	1209275	647938	561337	...	
72	862493	444639	417854	...	
184	725816	373794	352022	...	
7	589076	317917	271159	...	
119	418541	213084	205457	...	
274	300052	155475	144577	...	
449	531522	293208	238314	...	
380	324572	171152	153420	...	
225	378788	204320	174468	...	

	child_sex_ratio	effective_literacy_rate_total	\
185	902	90.28	
141	866	87.60	
72	940	89.59	
184	942	82.96	
7	853	89.62	
119	964	90.33	
274	930	87.14	
449	813	89.03	
380	896	91.61	
225	854	84.34	

```
     effective_literacy_rate_male  effective_literacy_rate_female  \
185                         93.32                           86.70
141                         91.44                           83.20
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
119                         93.47                           87.16
274                         89.08                           84.98
449                         92.76                           84.05
380                         95.13                           87.91
225                         90.61                           77.41

                  location  total_graduates  male_graduates  female_graduates  \
185        19.0760,72.8777          1802371          964964            837407
141        28.7041,77.1025          2221137         1210040           1011097
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
119  13.0826802,80.2707184           879695          487428            392267
274    22.572646,88.363895           818476          461615            356861
449  21.1702401,72.8310607           278795          160566            118229
380  18.5204303,73.8567437           656508          349022            307486
225  26.9124336,75.7872709           533148          319107            214041

       latitude   longitude
185     19.0760     72.8777
141     28.7041     77.1025
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
119  13.0826802  80.2707184
274   22.572646   88.363895
449  21.1702401  72.8310607
380  18.5204303  73.8567437
225  26.9124336  75.7872709

[10 rows x 24 columns]
```

In [37]:
```
# lets plot the top 10 literate cities on India map
plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()

lg=array(top10_female_kids_pop_cities['longitude'])
```

```
        lt=array(top10_female_kids_pop_cities['latitude'])
        pt=array(top10_female_kids_pop_cities['literates_total'])
        nc=array(top10_female_kids_pop_cities['name_of_city'])

        x, y = map(lg, lt)
        population_sizes_female_kids = top10_female_kids_pop_cities["literates_total"].apply(
        plt.scatter(x, y, s=population_sizes_female_kids, marker="o", c=population_sizes_femal


        for ncs, xpt, ypt in zip(nc, x, y):
            plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

        plt.title('Top 10 most literate Cities in India',fontsize=20)
```

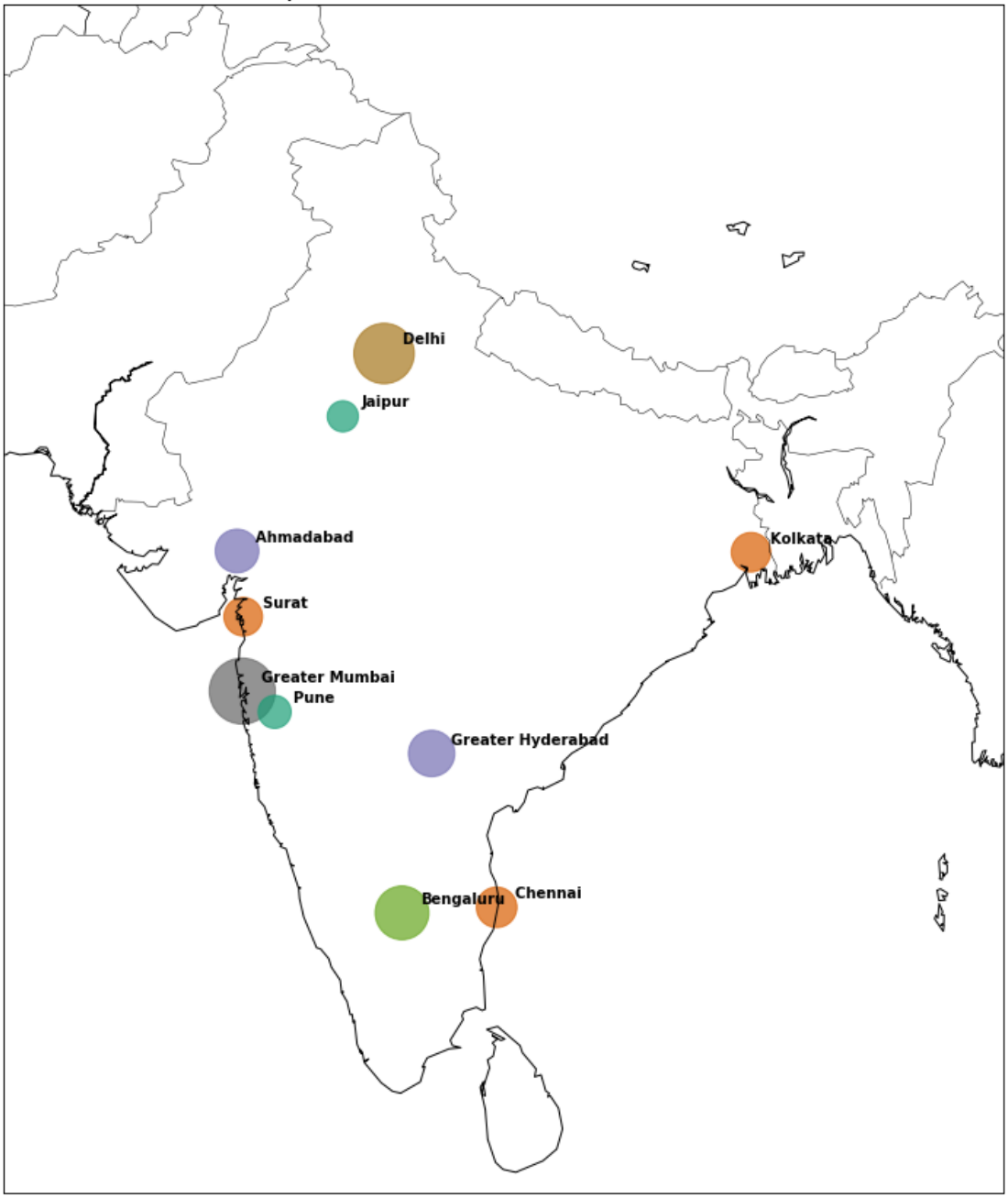
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

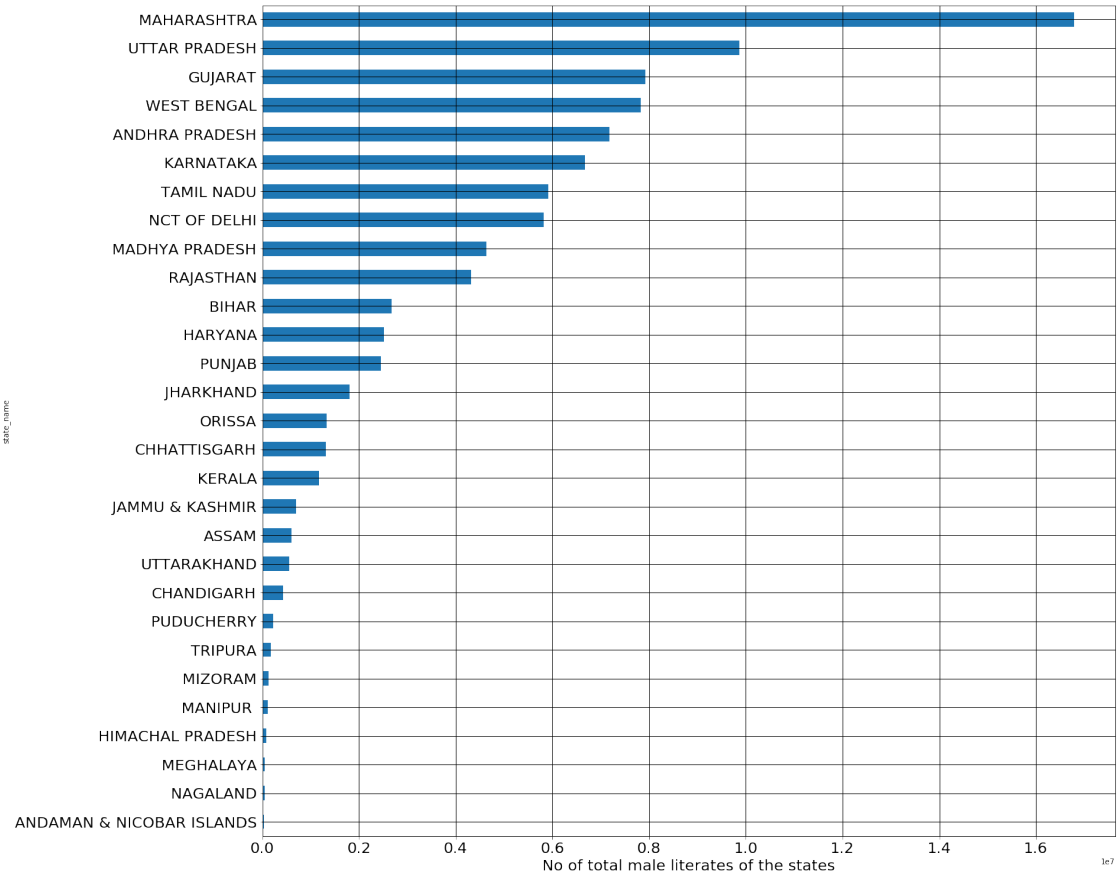
```
Out[37]: Text(0.5,1,'Top 10 most literate Cities in India')
```

1.20 Analysing Male Litracy rate of the states

```
In [38]: # # A bar chart to show the total male literates of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['literates_male'].sum().sort_values(ascending=T
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of total male literates of the states', fontsize = 20)
         plt.show ()
         # we can see again states like Maharashtra and UP have huge male literate population
```

Top 10 most literate Cities in India

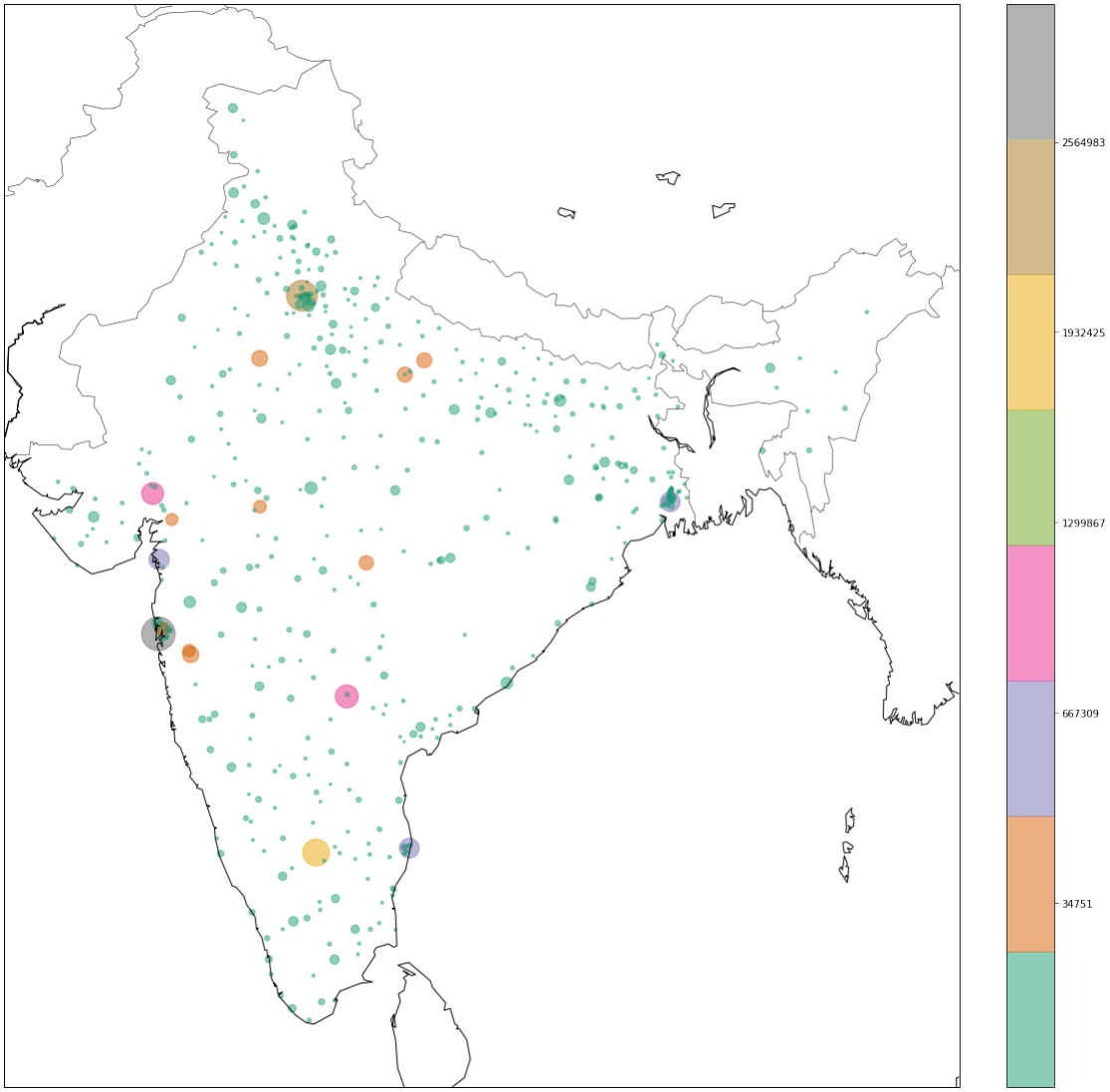




In [39]:
```
# Plotting the same on the map
population_sizes = cities["literates_male"].apply(lambda x: int(x / 5000))
colorbarValue = np.linspace(cities["literates_male"].min(), cities["literates_male"].m
                            num=10)
colorbarValue = colorbarValue.astype(int)

plot_map(population_sizes, colorbarValue)
# Major metro cities again shows higher male litracy rates
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e6416be80>
```



1.21 Top 10 cities where most of the male literates live

```
In [40]: # Lets find the top ten cities in which large number of males are literate
         print("The Top 10 Cities sorted according to the male literate Population (Descending
         top10_male_literate_cities = cities.sort_values(by='literates_male',ascending=False)
         top10_male_literate_cities=top10_male_literate_cities.head(10)
         top10_male_literate_cities
```

```
The Top 10 Cities sorted according to the male literate Population (Descending Order)
```

Out[40]:

	name_of_city	state_code	state_name	dist_code	\
185	Greater Mumbai	27	MAHARASHTRA	99	
141	Delhi	7	NCT OF DELHI	99	
72	Bengaluru	29	KARNATAKA	18	
184	Greater Hyderabad	28	ANDHRA PRADESH	99	
7	Ahmadabad	24	GUJARAT	7	
449	Surat	24	GUJARAT	25	
119	Chennai	33	TAMIL NADU	2	
274	Kolkata	19	WEST BENGAL	16	
380	Pune	27	MAHARASHTRA	25	
225	Jaipur	8	RAJASTHAN	12	

	population_total	population_male	population_female	\
185	12478447	6736815	5741632	
141	11007835	5871362	5136473	
72	8425970	4401299	4024671	
184	6809970	3500802	3309168	
7	5570585	2935869	2634716	
449	4462002	2538243	1923759	
119	4681087	2357633	2323454	
274	4486679	2362662	2124017	
380	3115431	1602137	1513294	
225	3073350	1619280	1454070	

	0-6_population_total	0-6_population_male	0-6_population_female	...	\
185	1139146	599007	540139	...	
141	1209275	647938	561337	...	
72	862493	444639	417854	...	
184	725816	373794	352022	...	
7	589076	317917	271159	...	
449	531522	293208	238314	...	
119	418541	213084	205457	...	
274	300052	155475	144577	...	
380	324572	171152	153420	...	
225	378788	204320	174468	...	

	child_sex_ratio	effective_literacy_rate_total	\
185	902	90.28	
141	866	87.60	
72	940	89.59	
184	942	82.96	
7	853	89.62	
449	813	89.03	
119	964	90.33	
274	930	87.14	
380	896	91.61	
225	854	84.34	

```
     effective_literacy_rate_male  effective_literacy_rate_female  \
185                         93.32                           86.70
141                         91.44                           83.20
72                          92.63                           86.25
184                         85.96                           79.79
7                           93.96                           84.81
449                         92.76                           84.05
119                         93.47                           87.16
274                         89.08                           84.98
380                         95.13                           87.91
225                         90.61                           77.41

                  location  total_graduates  male_graduates  female_graduates  \
185        19.0760,72.8777          1802371          964964            837407
141        28.7041,77.1025          2221137         1210040           1011097
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
449  21.1702401,72.8310607           278795          160566            118229
119  13.0826802,80.2707184           879695          487428            392267
274    22.572646,88.363895           818476          461615            356861
380  18.5204303,73.8567437           656508          349022            307486
225  26.9124336,75.7872709           533148          319107            214041

       latitude   longitude
185     19.0760     72.8777
141     28.7041     77.1025
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
449  21.1702401  72.8310607
119  13.0826802  80.2707184
274   22.572646   88.363895
380  18.5204303  73.8567437
225  26.9124336  75.7872709

[10 rows x 24 columns]
```

In [41]:
```
# Lets find the top ten cities in which large number of males are literate on the map

plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()
```

```
lg=array(top10_female_kids_pop_cities['longitude'])
lt=array(top10_female_kids_pop_cities['latitude'])
pt=array(top10_female_kids_pop_cities['literates_male'])
nc=array(top10_female_kids_pop_cities['name_of_city'])

x, y = map(lg, lt)
population_sizes_female_kids = top10_female_kids_pop_cities["literates_male"].apply(la
plt.scatter(x, y, s=population_sizes_female_kids, marker="o", c=population_sizes_femal


for ncs, xpt, ypt in zip(nc, x, y):
    plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

plt.title('Top 10 male litracy cities in India',fontsize=20)
```

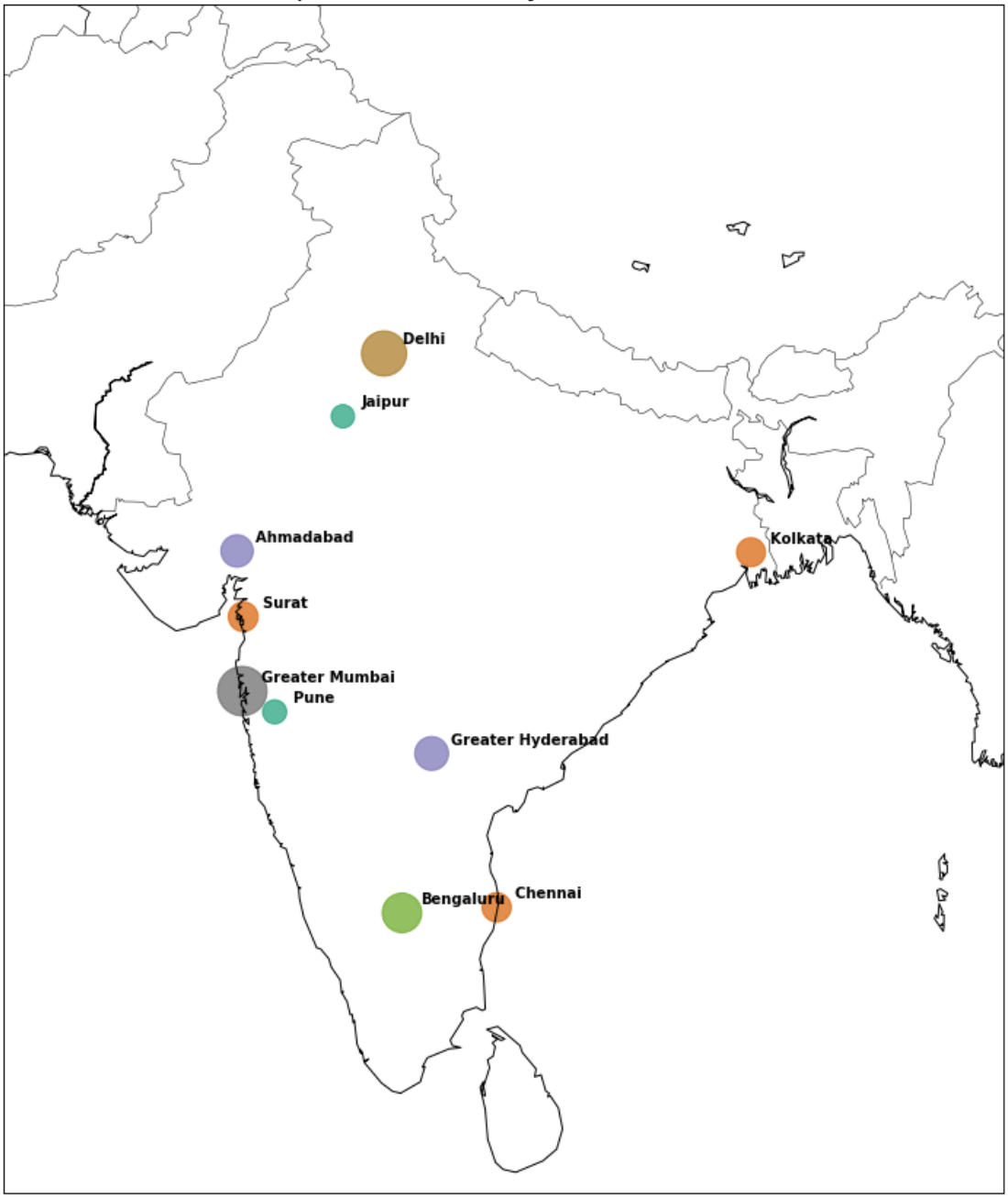
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

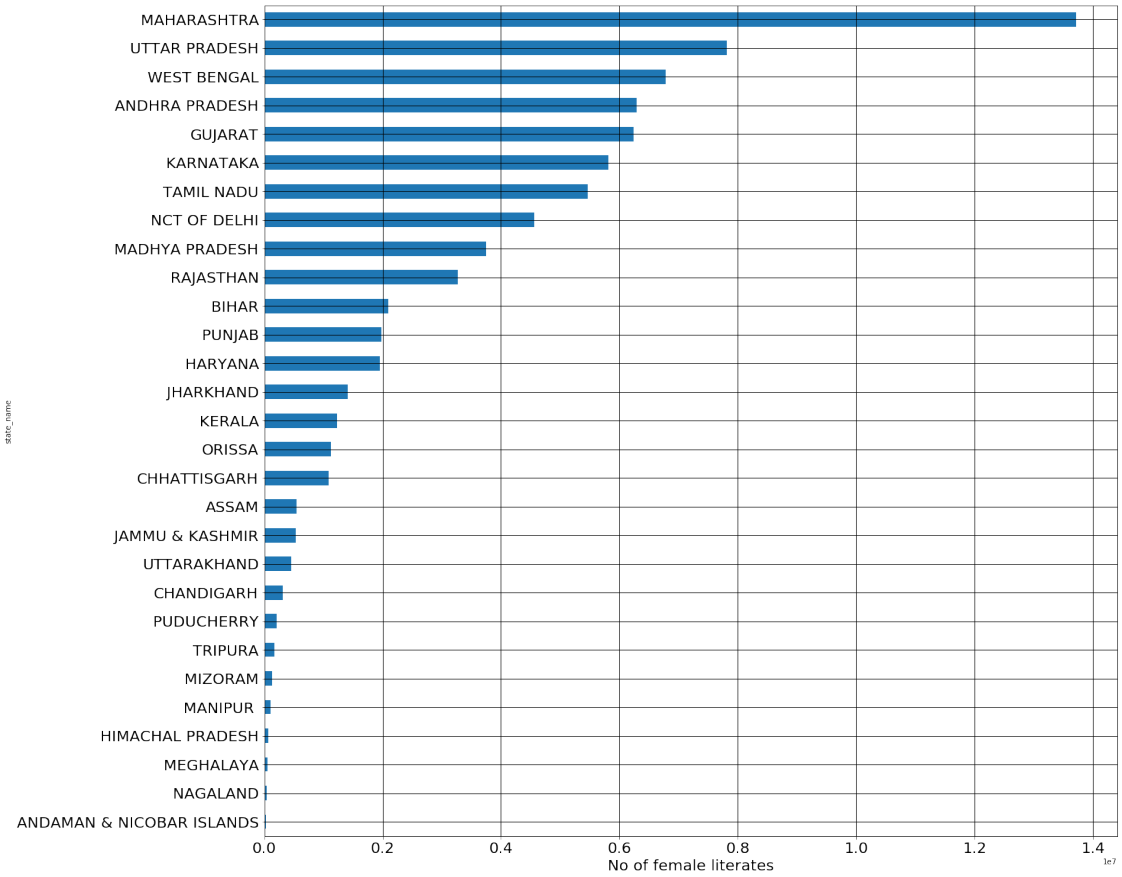
```
Out[41]: Text(0.5,1,'Top 10 male litracy cities in India')
```

1.22 Analysing Female Litracy rate of the states

```
In [42]: # A bar chart to show the female litracy population of the states
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['literates_female'].sum().sort_values(ascending=
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of female literates', fontsize = 20)
         plt.show ()
         # we can see again states like Maharashtra and UP have huge female literate populatio
```

Top 10 male litracy cities in India

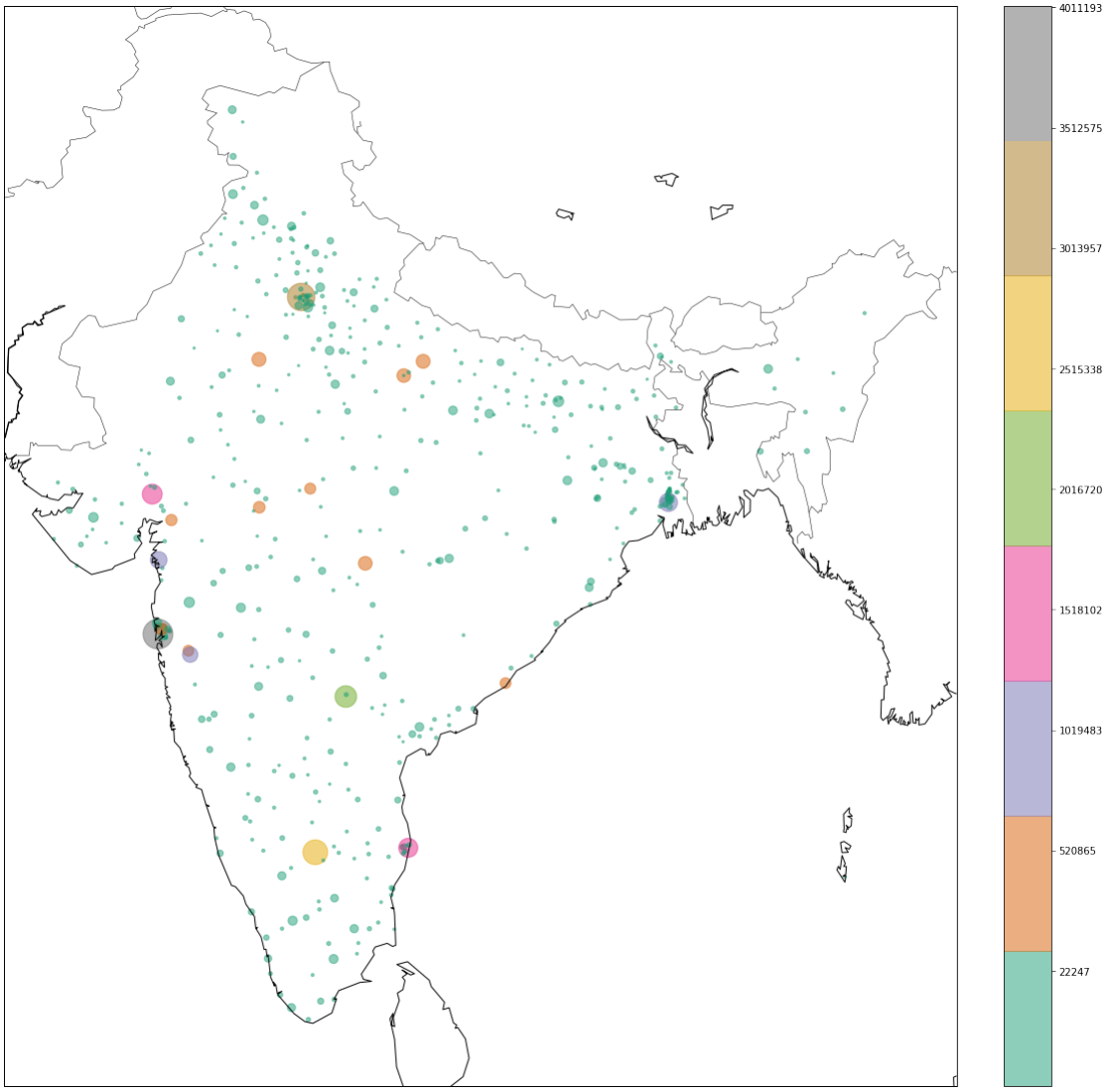





```
In [43]: # Plotting the same on the map
         population_sizes = cities["literates_female"].apply(lambda x: int(x / 5000))
         colorbarValue = np.linspace(cities["literates_female"].min(), cities["literates_female
                                     num=10)
         colorbarValue = colorbarValue.astype(int)

         plot_map(population_sizes, colorbarValue)
         # Major metro cities again shows higher female litracy rates
```

```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3236: MatplotlibDep
  b = ax.ishold()
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:3245: MatplotlibDep
    See the API Changes document (http://matplotlib.org/api/api_changes.html)
    for more details.
  ax.hold(b)
```

```
<matplotlib.figure.Figure at 0x25e62da8128>
```




```
In [44]: # Lets find the top ten cities in which large number of female literates live
         print("The Top 10 Cities sorted according to the Total Female literates Population (De
         top10_female_literates_cities = cities.sort_values(by='literates_female',ascending=Fal
         top10_female_literates_cities = top10_female_literates_cities.head(10)
         top10_female_literates_cities
```

```
The Top 10 Cities sorted according to the Total Female literates Population (Descending Order)
```

```
Out[44]:            name_of_city  state_code      state_name  dist_code  \
         185     Greater Mumbai           27     MAHARASHTRA         99
```

```
141              Delhi             7    NCT OF DELHI         99
72           Bengaluru            29       KARNATAKA         18
184  Greater Hyderabad            28  ANDHRA PRADESH         99
7            Ahmadabad            24         GUJARAT          7
119            Chennai            33      TAMIL NADU          2
274            Kolkata            19     WEST BENGAL         16
449              Surat            24         GUJARAT         25
380               Pune            27     MAHARASHTRA         25
225             Jaipur             8       RAJASTHAN         12

     population_total  population_male  population_female  \
185          12478447          6736815            5741632
141          11007835          5871362            5136473
72            8425970          4401299            4024671
184           6809970          3500802            3309168
7             5570585          2935869            2634716
119           4681087          2357633            2323454
274           4486679          2362662            2124017
449           4462002          2538243            1923759
380           3115431          1602137            1513294
225           3073350          1619280            1454070

     0-6_population_total  0-6_population_male  0-6_population_female  ...  \
185                1139146                599007                 540139  ...
141                1209275                647938                 561337  ...
72                  862493                444639                 417854  ...
184                 725816                373794                 352022  ...
7                   589076                317917                 271159  ...
119                 418541                213084                 205457  ...
274                 300052                155475                 144577  ...
449                 531522                293208                 238314  ...
380                 324572                171152                 153420  ...
225                 378788                204320                 174468  ...

     child_sex_ratio  effective_literacy_rate_total  \
185              902                          90.28
141              866                          87.60
72               940                          89.59
184              942                          82.96
7                853                          89.62
119              964                          90.33
274              930                          87.14
449              813                          89.03
380              896                          91.61
225              854                          84.34

     effective_literacy_rate_male  effective_literacy_rate_female  \
185                         93.32                           86.70
```

```
141                          91.44                          83.20
72                           92.63                          86.25
184                          85.96                          79.79
7                            93.96                          84.81
119                          93.47                          87.16
274                          89.08                          84.98
449                          92.76                          84.05
380                          95.13                          87.91
225                          90.61                          77.41

                  location  total_graduates  male_graduates  female_graduates  \
185        19.0760,72.8777          1802371          964964            837407
141        28.7041,77.1025          2221137         1210040           1011097
72         12.9716,77.5946          1591163          908363            682800
184        17.3850,78.4867          1164149          685402            478747
7     23.022505,72.5713621           769858          435267            334591
119  13.0826802,80.2707184           879695          487428            392267
274    22.572646,88.363895           818476          461615            356861
449  21.1702401,72.8310607           278795          160566            118229
380  18.5204303,73.8567437           656508          349022            307486
225  26.9124336,75.7872709           533148          319107            214041

       latitude   longitude
185     19.0760     72.8777
141     28.7041     77.1025
72      12.9716     77.5946
184     17.3850     78.4867
7     23.022505  72.5713621
119  13.0826802  80.2707184
274   22.572646   88.363895
449  21.1702401  72.8310607
380  18.5204303  73.8567437
225  26.9124336  75.7872709

[10 rows x 24 columns]
```

1.23 Top 10 cities in which most of the female literates live

In [45]:
```
# Lets find the top ten cities in which large number of female literates live

plt.subplots(figsize=(20, 15))
map = Basemap(width=1200000,height=900000,projection='lcc',resolution='l',
                    llcrnrlon=67,llcrnrlat=5,urcrnrlon=99,urcrnrlat=37,lat_0=28,lon_0=

map.drawmapboundary ()
map.drawcountries ()
map.drawcoastlines ()
```

```
        lg=array(top10_female_kids_pop_cities['longitude'])
        lt=array(top10_female_kids_pop_cities['latitude'])
        pt=array(top10_female_kids_pop_cities['literates_female'])
        nc=array(top10_female_kids_pop_cities['name_of_city'])

        x, y = map(lg, lt)
        population_sizes_female_kids = top10_female_kids_pop_cities["literates_female"].apply
        plt.scatter(x, y, s=population_sizes_female_kids, marker="o", c=population_sizes_femal


        for ncs, xpt, ypt in zip(nc, x, y):
            plt.text(xpt+60000, ypt+30000, ncs, fontsize=10, fontweight='bold')

        plt.title('Top 10 Female literates Populated Cities in India',fontsize=20)
```

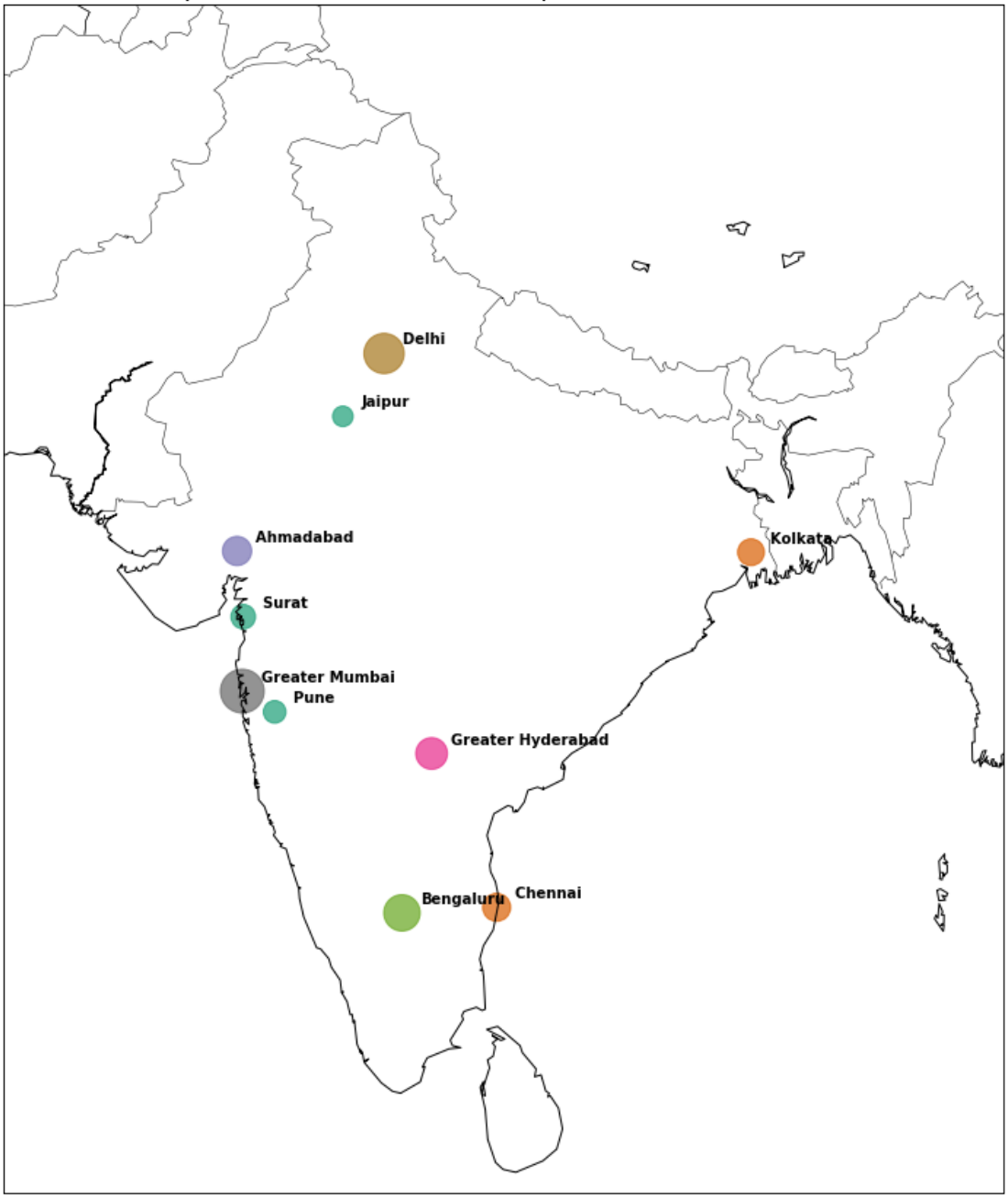
```
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1707: MatplotlibDep
  limb = ax.axesPatch
C:\ProgramData\Anaconda3\lib\site-packages\mpl_toolkits\basemap\__init__.py:1710: MatplotlibDep
  if limb is not ax.axesPatch:
```

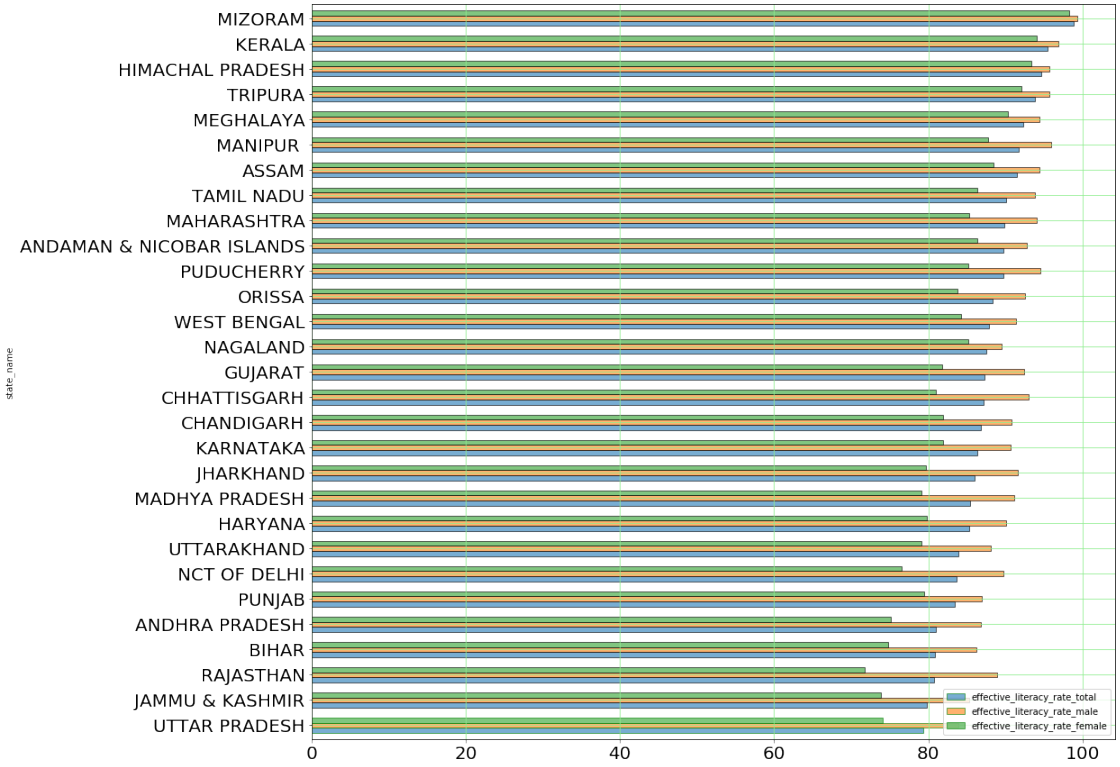
```
Out[45]: Text(0.5,1,'Top 10 Female literates Populated Cities in India')
```

1.24 Analyzing effective literacy rate

```
In [46]: # seperating effective literacy rate from the main dataset and sorting then in descen
         state_literacy_effective = cities[["state_name","effective_literacy_rate_total","effe


         state_literacy_effective.sort_values("effective_literacy_rate_total", ascending=True)
                               grid=True,
                               figsize=(16,15),
                               alpha = 0.6,
                               width=0.6,
                               stacked = False,
                               edgecolor="g",
                               fontsize = 20)
         plt.grid(b=True, which='both', color='lightGreen',linestyle='-')
         plt.show ()
         # from the below chart, Mizoram, Kerala and HP have highest effective literacy rate a
```

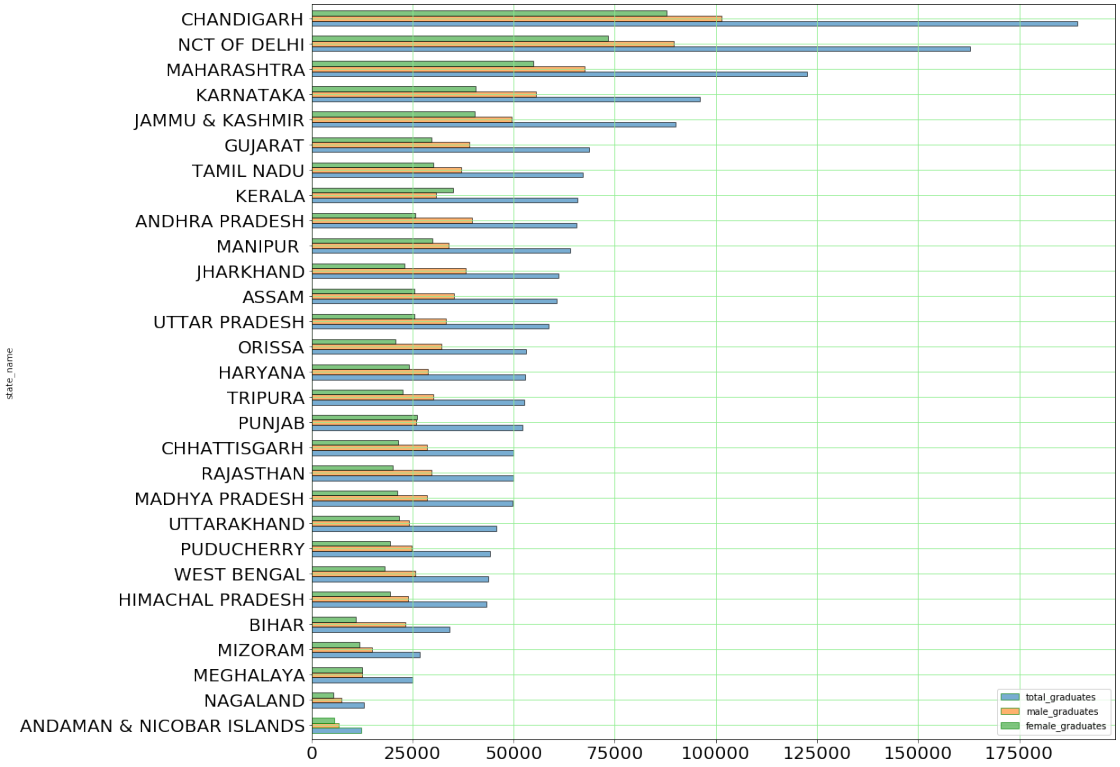
Top 10 Female literates Populated Cities in India





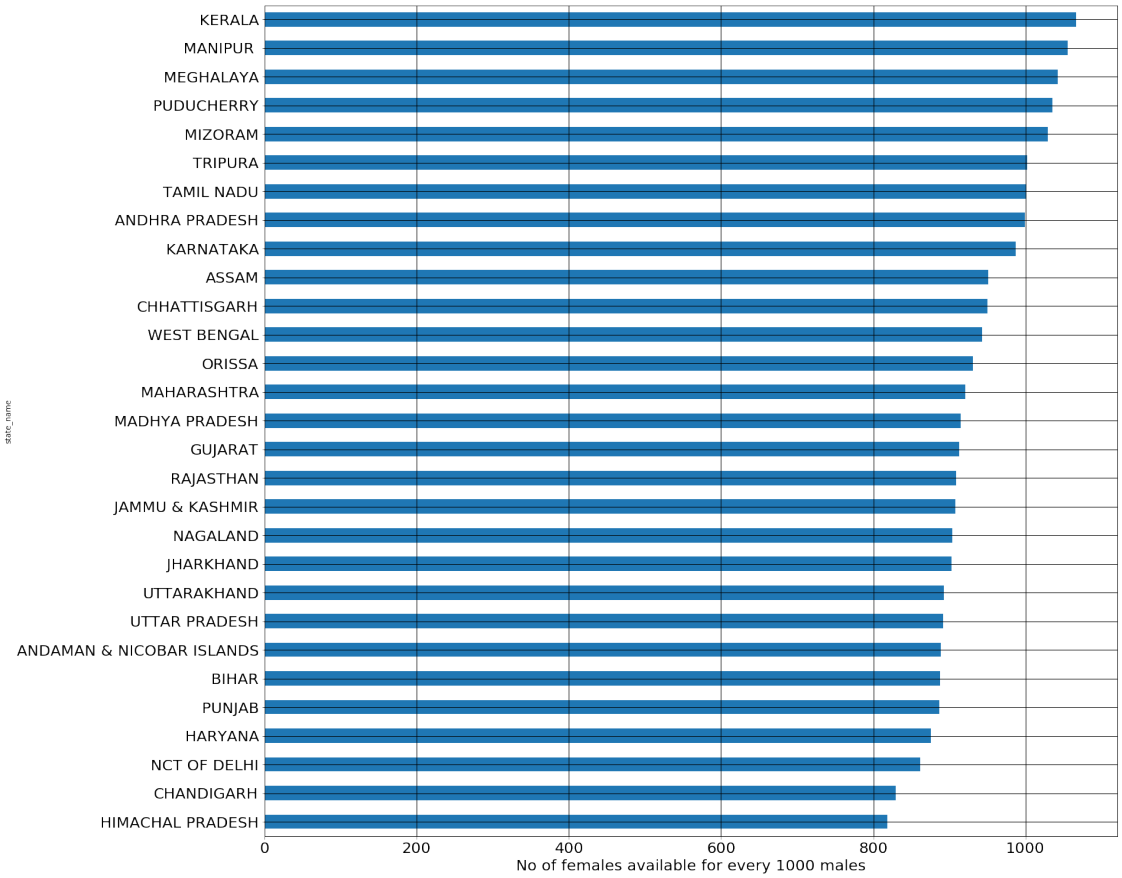
1.25 Analyzing graduates

```
In [47]: # seperating Graduates from the main dataset and sorting then in descending order
         state_graduates  = cities[["state_name",
                                    "total_graduates",
                                    "male_graduates",
                                    "female_graduates"]].groupby("state_name").agg({"to
                                                                                    "ma
                                                                                    "fe

         # Plotting the bar chart
         state_graduates.sort_values("total_graduates", ascending=True).plot(kind="barh",
                               grid=True,
                               figsize=(16,15),
                               alpha = 0.6,
                               width=0.6,
                               stacked = False,
                               edgecolor="g",
                               fontsize = 20)
         plt.grid(b=True, which='both', color='lightGreen',linestyle='-')
         plt.show ()
         # from the below Chandigarh, NCT of Delhi, Maharashta have most of their graduates li
         # we can note that Kerala and Meghalaya are the only states that have more number of
         # male graduates
```



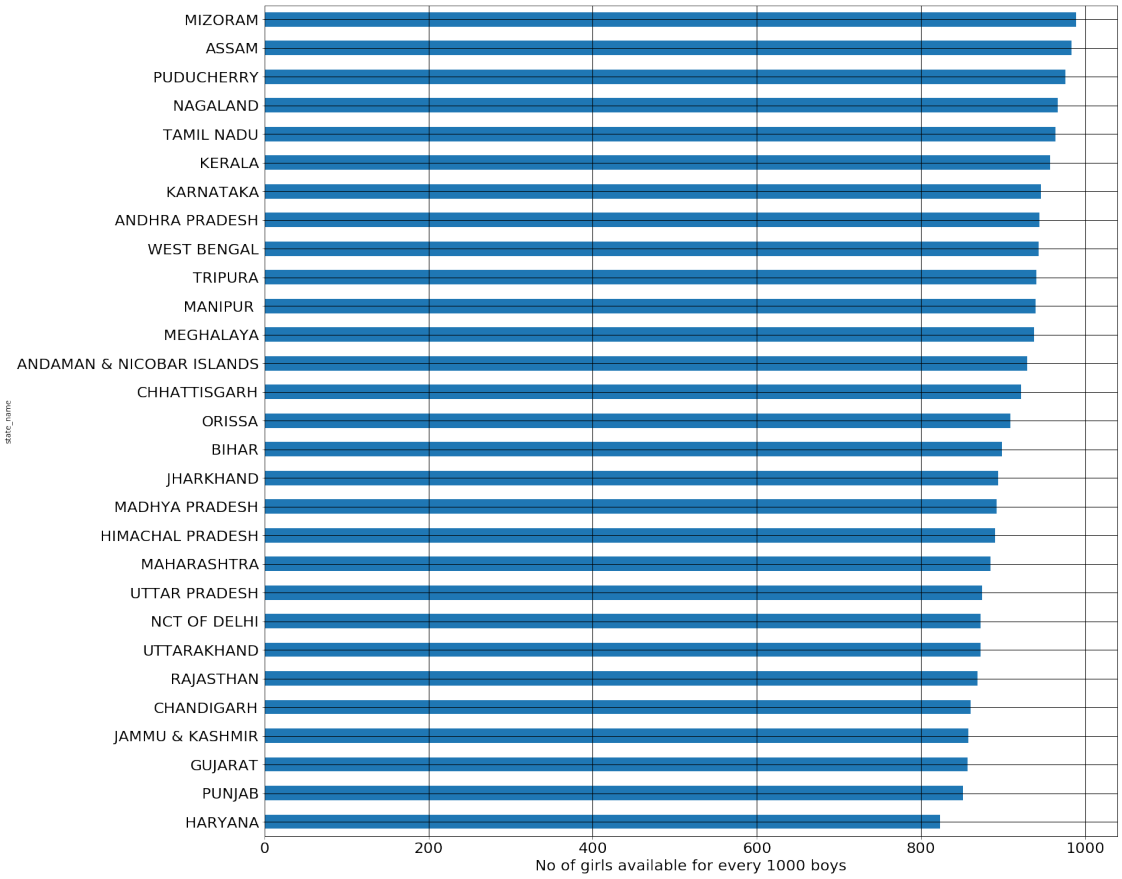
1.26 Analyzing Sex ratio across states

In [48]:
```
# A bar chart to show how many females are there for per 1000 males.
fig = plt.figure(figsize=(20,20))
states = cities.groupby('state_name')['sex_ratio'].mean().sort_values(ascending=True)
states.plot(kind="barh", fontsize = 20)
plt.grid(b=True, which='both', color='Black',linestyle='-')
plt.xlabel('No of females available for every 1000 males', fontsize = 20)
plt.show ()
# We can see that states of Kerala, Manipur, Meghalaya, Puducherry, Mizoram are havin
```



1.27 Analyzing Sex ratio across states for children below 6

```
In [49]: # A bar chart to show how many females are there for per 1000 males.
         fig = plt.figure(figsize=(20,20))
         states = cities.groupby('state_name')['child_sex_ratio'].mean().sort_values(ascending=
         states.plot(kind="barh", fontsize = 20)
         plt.grid(b=True, which='both', color='Black',linestyle='-')
         plt.xlabel('No of girls available for every 1000 boys', fontsize = 20)
         plt.show ()
         # Not even a single state have 1000 girls for every 1000 boys
```



1.28 Final Insights (Please remember that these insights are only for urban population. 70% of Indian population lives in Non-urban, Semi-urban and Village areas)

Uttar Pradesh and West Bengal are the two states from which most of the cities were taken, followed by Maharashtra and Andhra Pradesh.

Greater Mumbai, Delhi, Bengaluru are the top most populous cities of India in that order.

Maharashtra, Uttar Pradesh and Andhra Pradesh are the top 3 states, where most of the people lives in urban areas.

Maharashtra is the only state that have most of the male, female, male kids, female kids population living in urban areas.

Greater Mumbai is the most populated city in India (both men and women)

Delhi is the top most city that have high kids population

Maharashtra is the only state which have huge literate population living in urban areas.

Again Greater Mumbai have highest number of literates (both men and women), very closely followed by Delhi.

Mizoram is the state that have very high effective literacy rate, very closely followed by Kerala and Himachal Pradesh.

Not even a single state have more femal literates than male literates. Worst case is Rajasthan, where difference between effective literacy rate of men and women is very high.

Good thing is that almost all the states have effective literacy rate of more than 80 % (remeber this data is only for cities. 80% of Indian population lives in non-urban, semi-urban and village areas)

Interesting thing to note is, Kerala and Meghalaya are the only states where more female graduates are seen than male graduates in urban areas. Worst case is Bihar and Jharkand, where difference between men and women graduates is very high in urban areas itself.

This numbers might go up drastically in rural areas.

Sex ratio is defined as how many females are there for every 1000 males (atleast in India it is calculated in this way).

Kerala,Manipur, Meghalaya, Puduchery, Mizoram are the states where more than 1000 females are there for every > 1000 males. It means there are more females than males.

Worst case is Chandigarh and Himachal Pradesh where there are around 800 females for every 1000 males (which is clearly a bad sign)

When children below 6 are taken into account, not even a single state have 1000 girls for 1000 boys. This may change when these kids become adults, because age limit for this calculation is very narrow (6 yeras).