urbanization

July 24, 2023

1. datasets chosen : https://ourworldindata.org/urbanization
2. data preprocessing :
3. why do i chose the data :
   - the data collected spans from 1960 – 2021
   - part of the UNICEF-backed global World Development Indicator
   - high quality, accurate, clearly defined, and well-tested data
   - provides breakdown of data for each region, even country
4. the problem : Urbanization does not always mean positive, a lot of urban population can live in slums that negatively affect the economy.
5. Classify countries based on countries with high number of urbanization but low number of population living in slums
6. explain results with graphics, solution
7. remarks and suggestion
8. write article

1 Datasets Chosen

- Number of people living in urban areas

https://ourworldindata.org/urbanization

2 Data Preprocessing

- import dataset
- filter for year
- make another dataset for counting migration in 5 yrs period
- remove unnecessary columns

```python
import pandas as pd

# import all data
urban = pd.read_csv('./urban.csv')

urban.head()
```

```
         Entity Code  Year  Urban population  Rural population
0  Afghanistan  AFG  1960            724373         7898093.0
1  Afghanistan  AFG  1961            763336         8026804.0
```

```
2   Afghanistan  AFG  1962               805062          8163985.0
3   Afghanistan  AFG  1963               849446          8308019.0
4   Afghanistan  AFG  1964               896820          8458694.0
```

```
[ ]: slums = pd.read_csv('./slums.csv')
     slums.head()
```

```
[ ]:         Entity Code   Year  Urban population living in slums
     0  Afghanistan  AFG  2006                            3706745
     1  Afghanistan  AFG  2008                            3919555
     2  Afghanistan  AFG  2010                            4336525
     3  Afghanistan  AFG  2012                            4949920
     4  Afghanistan  AFG  2014                            5605681
```

```
[ ]: #exclude world, regions, and income groups
     urban = urban[urban['Entity'].isin([
         'Low-income countries',
         'Lower-middle-income countries',
         'Middle-income countries',
         'Upper-middle-income countries',
         'High-income countries',
         'Middle East and North Africa (WB)',
         'North America (WB)',
         'South Asia (WB)',
         'Sub-Saharan Africa (WB)',
         'Latin America and Caribbean (WB)',
         'Europe and Central Asia (WB)',
         'European Union (27)',
         'East Asia and Pacific (WB)',
         'World'])==False].reset_index(drop=True)

     urban.head()
```

```
[ ]:         Entity Code   Year  Urban population  Rural population
     0  Afghanistan  AFG  1960            724373         7898093.0
     1  Afghanistan  AFG  1961            763336         8026804.0
     2  Afghanistan  AFG  1962            805062         8163985.0
     3  Afghanistan  AFG  1963            849446         8308019.0
     4  Afghanistan  AFG  1964            896820         8458694.0
```

```
[ ]: #exclude world, regions, and income groups
     slums = slums[slums['Entity'].isin([
         'Low-income countries',
         'Lower-middle-income countries',
         'Middle-income countries',
         'Upper-middle-income countries',
         'High-income countries',
```

2

```
        'Middle East and North Africa (WB)',
        'North America (WB)',
        'South Asia (WB)',
        'Sub-Saharan Africa (WB)',
        'Latin America and Caribbean (WB)',
        'Europe and Central Asia (WB)',
        'European Union (27)',
        'East Asia and Pacific (WB)',
        'World'])==False].reset_index(drop=True)

slums.head()
```

```
[ ]:         Entity Code  Year  Urban population living in slums
     0  Afghanistan  AFG  2006                           3706745
     1  Afghanistan  AFG  2008                           3919555
     2  Afghanistan  AFG  2010                           4336525
     3  Afghanistan  AFG  2012                           4949920
     4  Afghanistan  AFG  2014                           5605681
```

```
[ ]: # filter for year, only 2020 is counted
     urban2 = urban.copy()
     urban2 = urban2[urban2['Year'].isin([2020])].reset_index(drop=True)

     # urban2.head()
     urban2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 215 entries, 0 to 214
Data columns (total 5 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Entity            215 non-null    object
 1   Code              213 non-null    object
 2   Year              215 non-null    int64
 3   Urban population  215 non-null    int64
 4   Rural population  215 non-null    float64
dtypes: float64(1), int64(2), object(2)
memory usage: 8.5+ KB
```

```
[ ]: slums2 = slums.copy()
     slums2 = slums2[slums2['Year'].isin([2020])].reset_index(drop=True)

     # slums2.head()
     slums2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 80 entries, 0 to 79
Data columns (total 4 columns):
```

```
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Entity                          80 non-null     object
 1   Code                            79 non-null     object
 2   Year                            80 non-null     int64
 3   Urban population living in slums 80 non-null    int64
dtypes: int64(2), object(2)
memory usage: 2.6+ KB
```

```python
# delete entity that doesnt exist in slums
urban3 = urban2.copy()
urban3 = urban3[urban3['Entity'].isin(slums2['Entity'])].reset_index(drop=True)
```

```python
# do an integrity check to make sure the entities are the same
urban3['Entity'].equals(slums2['Entity'])
```

True

```python
# delete unnecessary columns in slums
slums3 = slums2.copy()
slums3 = slums3.drop(columns=['Entity', 'Code', 'Year'])

slums3.head()
```

```
   Urban population living in slums
0                          7434756
1                            49354
2                          4249876
3                         13983955
4                           149117
```

```python
urban3 = urban3.drop(columns=['Code', 'Year'])

urban3.head()
```

```
        Entity  Urban population  Rural population
0  Afghanistan          10142913        28829316.0
1      Albania           1762645         1075204.0
2      Algeria          32038217        11413449.0
3       Angola          22338586        11089900.0
4      Armenia           1776315         1029293.0
```

```python
# combine urban and slums data
urban4 = pd.concat([urban3, slums3], axis=1, join='inner').
 ↪reset_index(drop=True)

urban4.head()
```

```
[ ]:          Entity  Urban population  Rural population  \
     0  Afghanistan          10142913        28829316.0
     1      Albania           1762645         1075204.0
     2      Algeria          32038217        11413449.0
     3       Angola          22338586        11089900.0
     4      Armenia           1776315         1029293.0

        Urban population living in slums
     0                           7434756
     1                             49354
     2                           4249876
     3                          13983955
     4                            149117
```

```
[ ]: # rename columns
     urban4 = urban4.rename(columns={'Entity':'Country', 'Urban population':
     ↪'Urban%', 'Rural population':'Rural%', 'Urban population living in slums':
     ↪'SlumInUrban%'})
     urban4.head()
```

```
[ ]:          Country     Urban%       Rural%  SlumInUrban%
     0  Afghanistan  10142913  28829316.0       7434756
     1      Albania   1762645   1075204.0         49354
     2      Algeria  32038217  11413449.0       4249876
     3       Angola  22338586  11089900.0      13983955
     4      Armenia   1776315   1029293.0        149117
```

```
[ ]: # normalize all values, convert to precentage
     urban5 = urban4.copy()
     urban5['SlumInUrban%'] = round((urban5['SlumInUrban%']/urban5['Urban%'])*100,2)
     urban5['Urban%'] = round((urban5['Urban%']/
     ↪(urban5['Urban%']+urban5['Rural%']))*100,2)

     # delete Rural% column
     urban5 = urban5.drop(columns=['Rural%'])

     urban5.head()
```

```
[ ]:          Country  Urban%  SlumInUrban%
     0  Afghanistan   26.03         73.30
     1      Albania   62.11          2.80
     2      Algeria   73.73         13.27
     3       Angola   66.83         62.60
     4      Armenia   63.31          8.39
```

3 Reason for choosing the data

- the data collected spans from 1960 – 2021
- part of the UNICEF-backed global World Development Indicator
- high quality, accurate, clearly defined, and well-tested data
- provides breakdown of data for each region, even country

4 The Problem

Behind the number of urbanization, lies a big problem that's arguably worse than low number of urbanization : the population of people living in urban slums. Urban slums is not a good place to live. It has strong ties to low income and high crime rate, which could cause other problems. Analyzing which countries has good urbanization number will tell us how that country deal with slums, which hopefully gives some insight on how to make urbanization better.

5 Classify Countries Urbanization

We will analyze the levels of urbanization in every country in the world in 2020. We will use k-means algorithm to cluster the data, and plotly.express to graph it

```python
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
import plotly.express as px

fig = px.scatter(urban5, x='Urban%', y='SlumInUrban%', color='Urban%',
hover_data=['Country','Urban%','SlumInUrban%'])
fig.show()
```

Here we see a nicely spread out data points on precentage of people live in slums variable against the urbanization precentage variable. This means each countries has their own unique challenges that results in very variate numbers. Next, lets classify the datapoints.

6 Explain the Results using graphs

```python
data = urban5.copy()

#change data to list of tuples
data = list(zip(data['Urban%'], data['SlumInUrban%']))

# fit kmeans model
kmeans = KMeans(n_init='auto' ,n_clusters=4)
kmeans.fit(data)

# make scatter plot using plotly
fig = px.scatter(urban5, x='Urban%', y='SlumInUrban%',
  ↪hover_data=['Country','Urban%','SlumInUrban%' ], color=kmeans.labels_)
```

```
fig.show()
```

The list of countries with their urbanization numbers are classified into 4 categories - high urban, low slum, bottom right corner - low urban, low slum, bottom left corner - high urban, high slum, center of screen - low urban, high slum, top left corner

7 Remarks and Suggestions

7.1 Data Availability Limitation

The number of countries shown in the graph is very low because of the limited data in population living in slums. Notable countries that are missing from the dataset includes: china, america, canada, italy, france, spain, germany, UK. As a result, the countries listed here are mostly developing/third world countries. The missing population living in slums data on big first world countries is very unfortunate. It could give us a great insight about how slums form in relatively advanced nations.

7.2 Ambiguity of Definition

Different countries have diverse criteria for defining an 'urban area', which includes variations in the metrics employed and the threshold population size. The minimum population threshold required for a settlement to be considered an 'urban area' depends on the data sources and can differ significantly from one country to another. For instance, Sweden and Denmark have a low threshold of merely 200 inhabitants, while Japan sets a considerably higher threshold of 50,000 inhabitants. It's worth noting that out of the total number of countries, 133 countries do not apply a minimum settlement population threshold in their definition of an 'urban area.' Instead, they may utilize alternative criteria such as population density, infrastructure development, pre-assigned city populations, or sometimes lack a distinct and explicit definition altogether. This variation in approaches further emphasizes the lack of a universal standard for categorizing urban areas across different nations.

Its the same case for the population living in urban slums dataset. However, UN-HABITAT offered a definition which states that
`a slum household as a group of individuals living under the same roof in an urban area who lack one or more of the following` : 1. Durable housing of a permanent nature that protects against extreme climate conditions.

2. Sufficient living space which means no more than three people sharing the same room.

3. Easy access to safe water in sufficient amounts at an affordable price.

4. Access to adequate sanitation in the form of a private or public toilet shared by a reasonable number of people.

Sometimes a fifth criterion is included:

5. Security of tenure that prevents forced evictions.

An important point to consider is that while a 'slum household' is assigned a single category, the actual conditions and level of deprivation can vary considerably among different slum households. Some households may lack only one of the mentioned criteria, while others might lack several of them. This variability highlights the complexity and diversity of living conditions within slums, and

a broad categorization may not fully capture the range of challenges and disparities experienced by individuals and families in these areas.

8 Slums, The Malice of Modern World Urbanization

8.1 Introduction

In the pursuit of progress and development, the world has witnessed an unprecedented wave of urbanization. With cities evolving into hubs of economic and social activities, they magnetically draw people seeking better opportunities and improved living standards. However, this rapid urban growth has also given rise to a dark underbelly - slums. These pockets of poverty and deprivation stand as stark reminders of the challenges that accompany modern world urbanization. In this article, we delve into the malice of slums, exploring their causes, impact, and potential solutions.

8.2 Urbanization and the Proliferation of Slums

Urbanization, the process of population shift from rural to urban areas, lies at the heart of the slum phenomenon. As cities grow exponentially, they attract migrants seeking better economic prospects, education, and improved living standards. While urbanization can be a powerful driver of economic growth and cultural exchange, the pace and scale of this phenomenon often outstrip the capacity of cities to accommodate their growing populations adequately.

Rapid urbanization presents a double-edged sword for governments and urban planners. On one hand, it offers the potential for increased economic opportunities and improved infrastructure, while on the other, it gives rise to a surge in informal settlements – the slums. The pressing need for affordable housing, coupled with inadequate urban planning and insufficient resources, forces many new migrants and low-income residents to settle in substandard living conditions.

Slums, in essence, are the manifestation of a city's struggle to cope with the challenges posed by urbanization. Without adequate provisions for affordable housing and social amenities, slums emerge as a result of a demand-supply mismatch. As urbanization continues to reshape our world, addressing the issue of slums becomes an urgent imperative.

8.3 Social and Economic Implications:

The existence of slums poses significant social and economic challenges for societies. Concentrated poverty and overcrowding foster crime, social unrest, and increased vulnerability to exploitation. Slum dwellers often find themselves trapped in a cycle of informality, unable to access formal employment or credit, which perpetuates economic marginalization. Furthermore, slums strain urban infrastructure and services, hampering overall city development and exacerbating inequality.

8.4 Classification of Countries's Urbanization Data

Based on our findings, we classified countries in our dataset into 4 categories: 1. high urban, low slum, bottom right corner 2. low urban, low slum, bottom left corner 3. high urban, high slum, center of screen 4. low urban, high slum, top left corner

We'll look into a country in each category and analyze their methods on handling urbanization and slums.

8.4.1 High Urban, Low Slum - Ireland

Ireland is a sub-tropical island country located in the north atlantic ocean, north western europe. The geography of ireland is made of low mountains with vast plains covering the landscape. Several rivers extending inland gives Ireland great advantage in agriculture.

Ireland's culture is a vibrant tapestry steeped in history and folklore. Traditional music, dance, and literature hold a significant place, with renowned writers like James Joyce and W.B. Yeats contributing to its literary heritage. Festivals like St. Patrick's Day celebrate the country's spirit, and the Gaelic language adds to its identity. With a strong religious heritage and modern expressions of art and creativity, Ireland's culture reflects its deep-rooted traditions and dynamic present.

Despite having very strong agriculture background thats still running in the present, Ireland manages to reach an impressive number of urbanization and low number of population living in slums. Ireland has a diverse and robust economy, with key sectors including pharmaceuticals, technology, finance, and agriculture. Dublin, Ireland's capital is home to one of the biggest company in earth: Google. Ireland also attracted several other high-profile companies including Facebook, Apple, Pfizer, J&J, Citibank, and JP-Morgan-Chase. These companies gave strong incentive to urbanize to big cities like Dublin, Belfast, and Cork.

Ireland has handled urbanization through sustainable urban planning, regional development, and affordable housing initiatives. Investment in infrastructure and utilities supports urban growth, while urban regeneration revitalizes older areas. Emphasizing community engagement, the country preserves cultural heritage and focuses on balanced development to avoid over-reliance on major cities. Ireland continues to adapt strategies to ensure a sustainable and well-managed urbanization process, promoting a resilient and inclusive future.

8.4.2 Low Urban, Low Slum - Vietnam

Vietnam, officially known as the Socialist Republic of Vietnam, is a Southeast Asian country located on the eastern coast of the Indochinese Peninsula. It shares borders with China to the north, Laos to the northwest, and Cambodia to the southwest. With a population of over 95 million people, Vietnam is the 15th most populous country in the world. Its capital city is Hanoi, while Ho Chi Minh City is the largest city and a vital economic hub.

Vietnam's history is characterized by a rich tapestry of ancient civilizations, colonial influence, and wars, including the Vietnam War, which ended in 1975. Today, Vietnam is a rapidly developing nation with a dynamic economy driven by agriculture, manufacturing, and services sectors. The country's natural beauty, cultural diversity, and delicious cuisine make it an increasingly popular destination for tourism. Vietnam is known for its resilient people, strong sense of community, and ongoing efforts to balance traditional values with modernization and progress.

Vietnam, an agriculturally focused country, stands as one of the largest exporters of products like rice, coffee, seafood, and fruits. Consequently, a significant portion of its population resides in rural areas, engaged in farming and cultivation. This agricultural emphasis has contributed to a relatively low percentage of just 5.7% of the population living in slums, as the rural lifestyle holds substantial value compared to anything else. Hence, there is limited pressure or incentive for rural residents to migrate to urban areas.

8.4.3 High Urban, High Slum, - Philippines

The Philippines, officially known as the Republic of the Philippines, is an archipelagic country located in Southeast Asia. It comprises over 7,600 islands, with Luzon, Visayas, and Mindanao being the three main island groups. The capital city is Manila. With a population of approximately 110 million people, the Philippines is the 13th most populous country in the world.

The nation's history is marked by a blend of indigenous, Spanish, American, and Asian influences. Filipino culture is known for its warm hospitality, rich traditions, and vibrant festivals. The economy of the Philippines is diverse, with sectors such as agriculture, manufacturing, services, and tourism contributing to its growth. The country is famous for its stunning natural beauty, including pristine beaches, lush landscapes, and active volcanoes. However, the Philippines faces challenges related to poverty, inequality, and natural disasters. Despite these hurdles, the Filipino people exhibit resilience, strong community ties, and a spirit of optimism for the future.

Much like Vietnam, the Philippines was primarily an agricultural country, but it successfully attracted foreign investors to engage in business. The nation's ample unoccupied lands, large population, and comparatively lower education and skills level created an ideal environment for offshoring manufacturing plants, particularly in garments, footwear, electronics, and automobiles. The influx of foreign investment into urban areas became a golden opportunity for rural residents seeking better pay, housing, and overall quality of life, prompting significant rural-to-urban migration.

However, despite this urbanization boom, challenges arise due to insufficient government action, intense competition, and limited job opportunities, leaving many newly-urbanized individuals struggling to secure stable employment. Fearful of returning to their rural homes seen as a failure, they opt to create makeshift settlements on the city outskirts, adapting to their circumstances and making the best of life.

8.4.4 Low Urban, High Slum, - Afghanistan

Afghanistan is a landlocked country located in Central Asia, sharing borders with Pakistan, Iran, Turkmenistan, Uzbekistan, Tajikistan, and China. With a diverse landscape, it encompasses mountains, deserts, and fertile valleys. Kabul serves as its capital and largest city. Afghanistan's history is marked by a mix of ancient civilizations and a complex geopolitical past. Afghanistan's geography is diverse, characterized by mountains, plateaus, deserts, and fertile valleys. The Hindu Kush mountain range runs through the country, forming a rugged and challenging terrain. The country experiences a mix of continental and arid climates, with hot summers and cold winters in most regions.

Afghanistan's populace is predominantly Muslim, with Islam being the dominant and official religion of the country. The vast majority of Afghans are followers of Sunni Islam, specifically belonging to the Hanafi school of thought, which is one of the four major Sunni schools of jurisprudence. Sunni Islam in Afghanistan is influenced by traditional practices and customs that have evolved over centuries. There are also small communities of Shia Muslims and other religious minorities in the country. Religion holds a central place in Afghan society, influencing various aspects of daily life, cultural norms, and social interactions.

The main cause low number of urbanization and high population living in slum is armed conflict.The armed conflict in Afghanistan is a complex and protracted conflict that has spanned several decades. It originated in the late 1970s when the Soviet Union invaded Afghanistan, leading to a decade-long war known as the Soviet-Afghan War. Following the Soviet withdrawal in 1989, internal conflicts

between various Afghan factions ensued, culminating in the rise of the Taliban, an extremist Islamist group, which took control of most of the country in the mid-1990s. The U.S.-led invasion of Afghanistan in 2001 aimed to dismantle terrorist networks, remove the Taliban from power, and promote stability. Despite efforts to establish a democratic government, Afghanistan has faced ongoing violence and insurgency from the Taliban and other militant groups.

Those who live in rural areas want to stay in rural areas to avoid conflicts that are happening in big cities. Meanwhile those who lives in the city often find their houses get destroyed due to gunfire and bombing, so they resort to living in makeshift houses in safer areas.

8.5   Conclusion:

The phenomenon of slums amidst rapid urbanization serves as a poignant reminder of the challenges faced by cities in the 21st century. Urbanization presents immense opportunities for progress, innovation, and cultural exchange. However, its hasty and unchecked pace can lead to the growth of informal settlements, where millions endure dire living conditions.

To address the malice of slums, a concerted effort is required from governments, urban planners, and society at large. Sustainable urban planning, affordable housing initiatives, and empowering slum dwellers are crucial steps towards fostering inclusive and vibrant cities. By recognizing the intricate relationship between urbanization and slums and adopting proactive measures, we can aspire to build cities that are beacons of hope, prosperity, and compassion for all their inhabitants. Only then can we truly embrace the transformative potential of urbanization while leaving no one behind.

9   Sources

https://ourworldindata.org/urbanization#definitions-measurement https://www.geeksforgeeks.org/ml-classification-vs-clustering/ https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html https://www.w3schools.com/python/python_ml_k-means.asp