Визуализация данных это наиболее эффективный метод иллюстрации и объяснения сложной информации, особенно числовых данных, в простой и удобоваримой форме. Кроме того, при грамотном исполнении, инфографика при интерпретации данных позволяет уменьшить или смягчить систематические ошибки. Один из заслуживающих особого внимания видов визуализации — анимированные фоновые картограммы или хороплеты. И в данной статье мы, используя эти виды, визуализируем пандемию коронавируса, что на данный момент очень актуально.

Но прежде чем начать, давайте разберем определения. Фоновая картограмма или хороплет — это тип тематической карты, на которой области или регионы заштрихованы пропорционально и в соответствии со значением заданного измерения данных.

Статические хороплеты наиболее полезны, когда надо сравнить какие‑то показатели по регионам. Например, сравнить уровень преступности в данный момент в каждом субъекте Российской Федерации, что можно визуализировать с помощью статической фоновой картограммы.

Анимированная или динамическая фоновая картограмма похожа на статическую, за исключением того, что добавляется время. Это третье измерение делает визуализацию исключительно интересной и мощной.

Для визуализации пандемии коронавируса, я использовал набор данных Novel Corona Virus 2019 от Kaggle, который можно найти [здесь](http://waksoft.susu.ru/53644d/CRFMEUtYSklER0MaXQIDAVwAS1VbDxZHTFcHWFFeRAALDk0MWRBKCFxGUVgbAAsUXwsEG0ILS0FKHlQEAQ4bBQARWRJdFg==/). Набор данных получен из нескольких источников, включая Всемирную организацию здравоохранения, Национальную комиссию здравоохранения Китайской Народной Республики и Центры США по контролю за заболеваниями.

**Статический хороплет**

Начнем мы с создания статического хороплета. Работать я буду в привычном мне PyCharm. Первым делом, я открываю редактор и создаю проект. После создания, я добавляю в него папку, под названием “data”, где у меня будет хранится файл csv формата, скаченный с Kaggle и содержащий данные о пандемии по всему миру.

После этого, я, через терминал, устанавливаю необходимые библиотеки. Команды для установки представлены ниже:

pip install numpy

pip install pandas

pip install plotly

Проделав операции выше, создаем python файл, и мы готовы к визуализации. Код для этой инфографики выглядит следующим образом:

#Статический хороплет  
  
  
# Импорт библиотек  
import numpy as np  
import pandas as pd  
import plotly as py  
import plotly.express as px  
import plotly.graph\_objs as go  
from plotly.subplots import make\_subplots  
from plotly.offline import download\_plotlyjs, init\_notebook\_mode, plot, iplot  
  
# Чтение данных  
df = pd.read\_csv("data/covid\_19\_data.csv")  
  
# Переименуем колонки  
df = df.rename(columns={'Country/Region':'Страна'})  
df = df.rename(columns={'ObservationDate':'Дата'})  
df = df.rename(columns={'Confirmed':'Подтверждено'})  
  
# Манипуляции с оригиналом Dataframe  
df\_countries = df.groupby(['Страна', 'Дата']).sum().reset\_index().sort\_values('Дата', ascending=False)  
df\_countries = df\_countries.drop\_duplicates(subset = ['Страна'])  
df\_countries = df\_countries[df\_countries['Подтверждено']>0]  
  
# Создание фоновой картограммы  
fig = go.Figure(data=go.Choropleth(  
 locations = df\_countries['Страна'],  
 locationmode = 'country names',  
 z = df\_countries['Подтверждено'],  
 colorscale = 'Reds',  
 marker\_line\_color = 'black',  
 marker\_line\_width = 0.5,  
))  
fig.update\_layout(  
 title\_text = 'Подтверждённые заболевания 28 марта 2020',  
 title\_x = 0.5,  
 geo=dict(  
 showframe = False,  
 showcoastlines = False,  
 projection\_type = 'equirectangular'  
 )  
)  
  
fig.write\_html('Static\_choroplete.html', auto\_open=True)

Здесь, мы импортируем установленные ранее библиотеки, после читаем csv файл из которого будем брать данные, указывая при этом его местонахождение и название. Далее, переименовываем колонки для удобного восприятия и проводим некоторые манипуляции с оригиналом Dataframe: группировки и возвращения ряда с удаленными повторяющимися значениями. После создаем фоновую картограмму и последней командой мы указываем на создание и запуск html файла, содержащий наши данные. Результат представлен ниже:

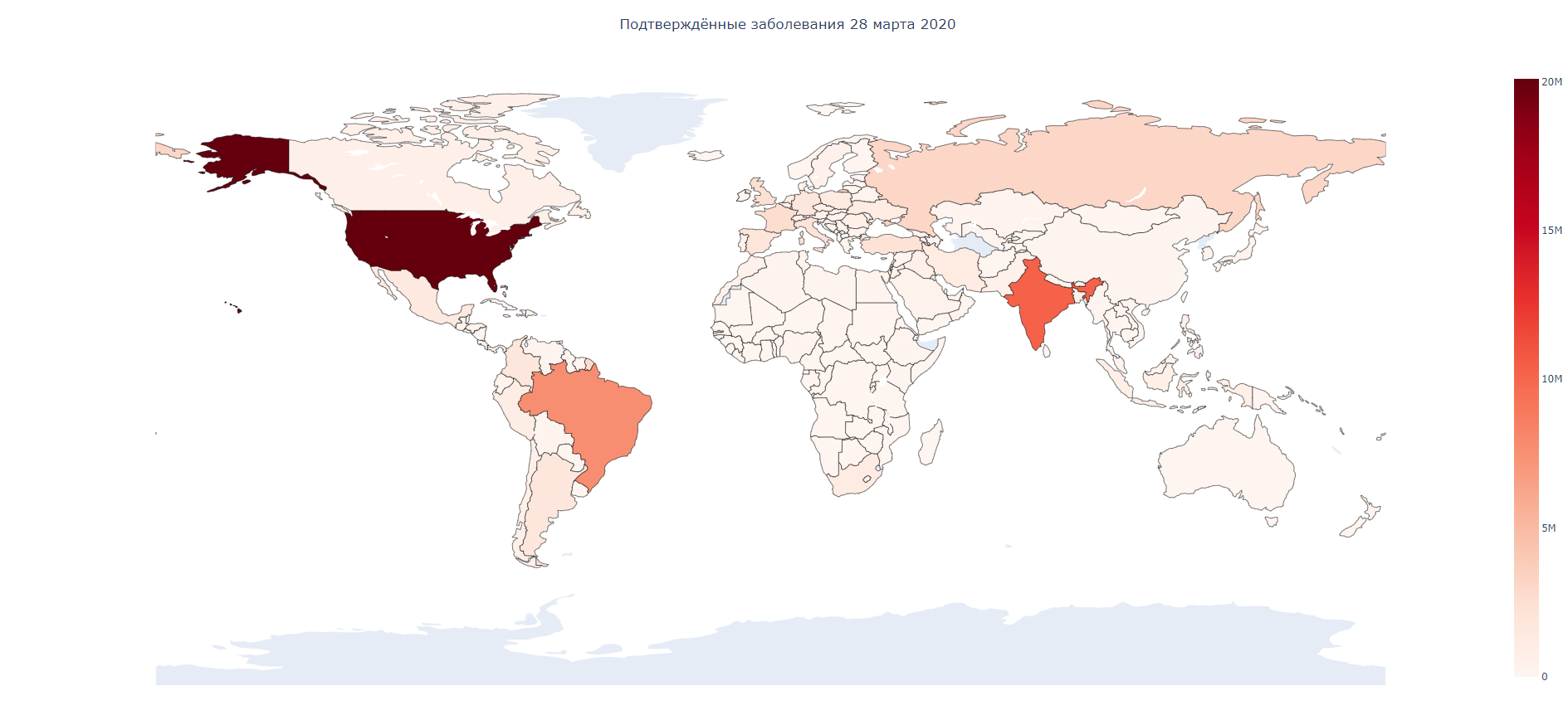


Рисунок 1 - Статическая фоновая картограмма

На рисунке, представлена ​​статическая диаграмма общего количества подтвержденных случаев коронавируса по странам по состоянию на 28 марта 2020 года. Вы можете видеть, что страны с наибольшим количеством подтверждённых случаев заболевания — это США, Китай и Италия, а также несколько других европейских стран.

### Анимированная фоновая картограмма

Сейчас, вы увидите, насколько, по сравнению со статической, более эффектна и интересна анимированная фоновая картограмма. Для этого, создаем новый python файл и прописываем следующий код:

#Анимированная фоновая картограмма  
  
# Импорт библиотек  
import numpy as np  
import pandas as pd  
import plotly as py  
import plotly.express as px  
import plotly.graph\_objs as go  
from plotly.subplots import make\_subplots  
from plotly.offline import download\_plotlyjs, init\_notebook\_mode, plot, iplot  
  
# Чтение данных  
df = pd.read\_csv("data/covid\_19\_data.csv")  
  
# Переименуем колонки  
df = df.rename(columns={'Country/Region': 'Страна'})  
df = df.rename(columns={'ObservationDate': 'Дата'})  
df = df.rename(columns={'Confirmed': 'Подтверждено'})  
  
# Манипуляции с оригиналом Dataframe  
df\_countrydate = df[df['Подтверждено'] > 0]  
df\_countrydate = df\_countrydate.groupby(['Дата', 'Страна']).sum().reset\_index()  
df\_countrydate  
  
# Создание фоновой картограммы  
fig = px.choropleth(df\_countrydate,  
 locations="Страна",  
 locationmode="country names",  
 color="Подтверждено",  
 hover\_name="Страна",  
 animation\_frame="Дата"  
 )  
fig.update\_layout(  
 title\_text='Глобальное распространение короновируса',  
 title\_x=0.5,  
 geo=dict(  
 showframe=False,  
 showcoastlines=False,  
 ))  
  
fig.write\_html('Animated\_background\_cartogram.html', auto\_open=True)

В данном коде, мы проделываем те же операции, что и при статическом хороплете, но с небольшими отличиями. Результат представлен ниже:

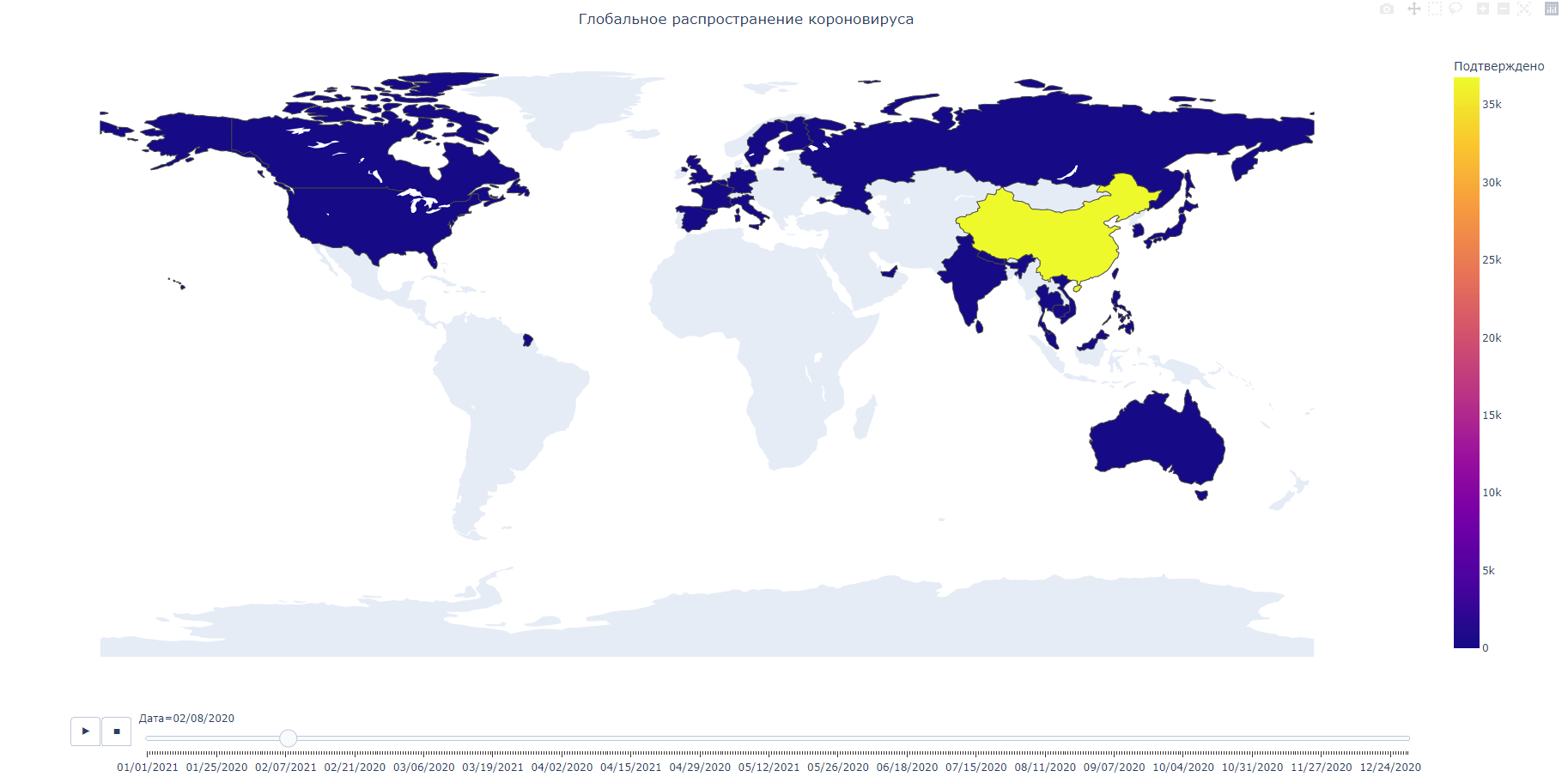


Рисунок 2 – Анимированная фоновая картограмма

На рисунке 2, мы смотрим на общее количество подтвержденных случаев коронавируса по странам во времени. И вы можете видеть, что до недавнего времени больше всего заболевших было в Китае.