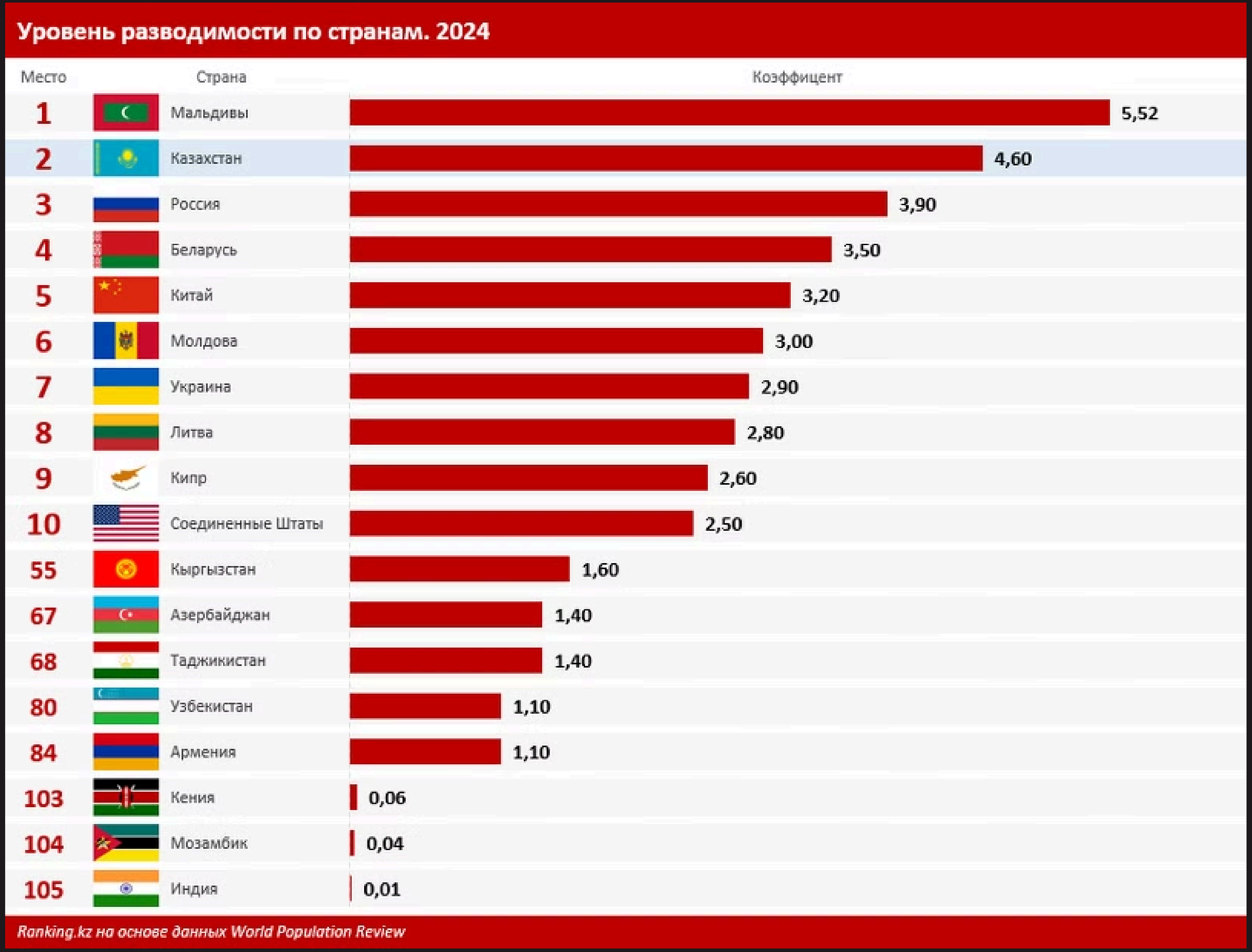
Predicting Divorce Rates in Kazakhstan

Understanding the problem



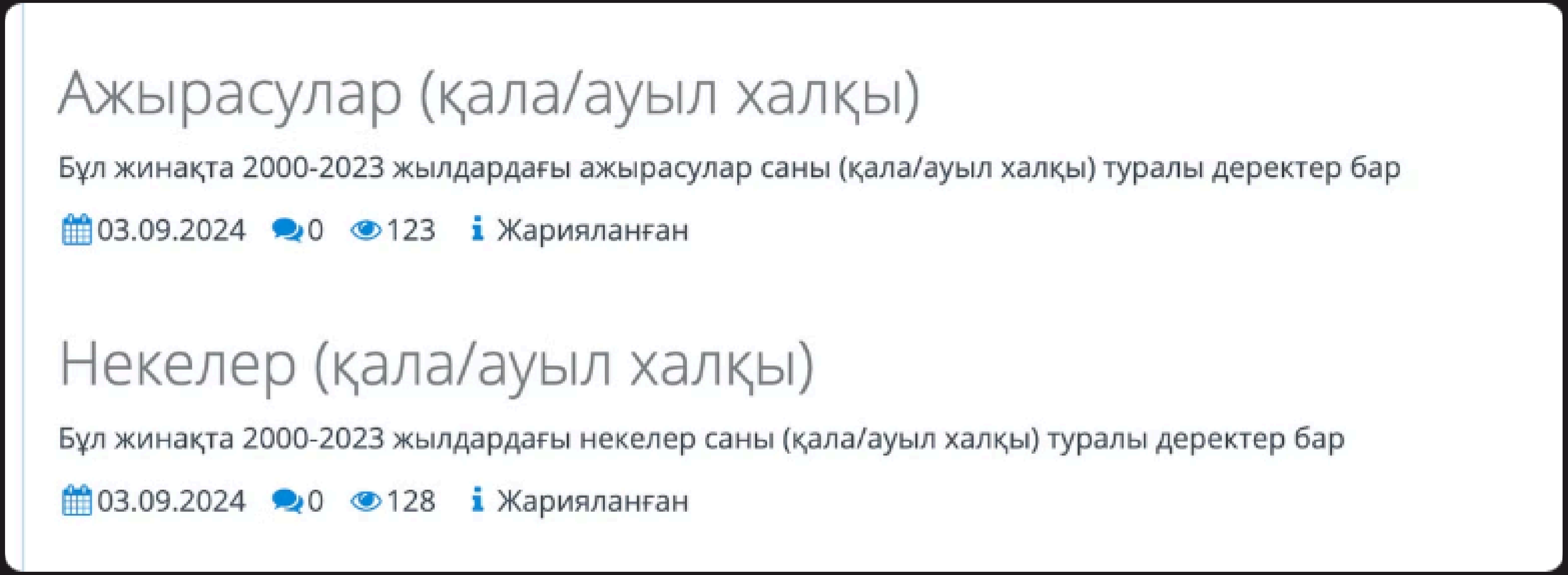
Уровень разводимости по странам. 2024

Место	Страна	Коэффициент
1	Мальдивы	5,52
2	Казахстан	4,60
3	Россия	3,90
4	Беларусь	3,50
5	Китай	3,20
6	Молдова	3,00
7	Украина	2,90
8	Литва	2,80
9	Кипр	2,60
10	Соединенные Штаты	2,50
55	Кыргызстан	1,60
67	Азербайджан	1,40
68	Таджикистан	1,40
80	Узбекистан	1,10
84	Армения	1,10
103	Кения	0,06
104	Мозамбик	0,04
105	Индия	0,01

Ranking.kz на основе данных World Population Review

Data collection

data.egov.kz



Ажырасулар (қала/ауыл халқы)

Бұл жинақта 2000-2023 жылдардағы ажырасулар саны (қала/ауыл халқы) туралы деректер бар

03.09.2024 0 123 Жарияланған

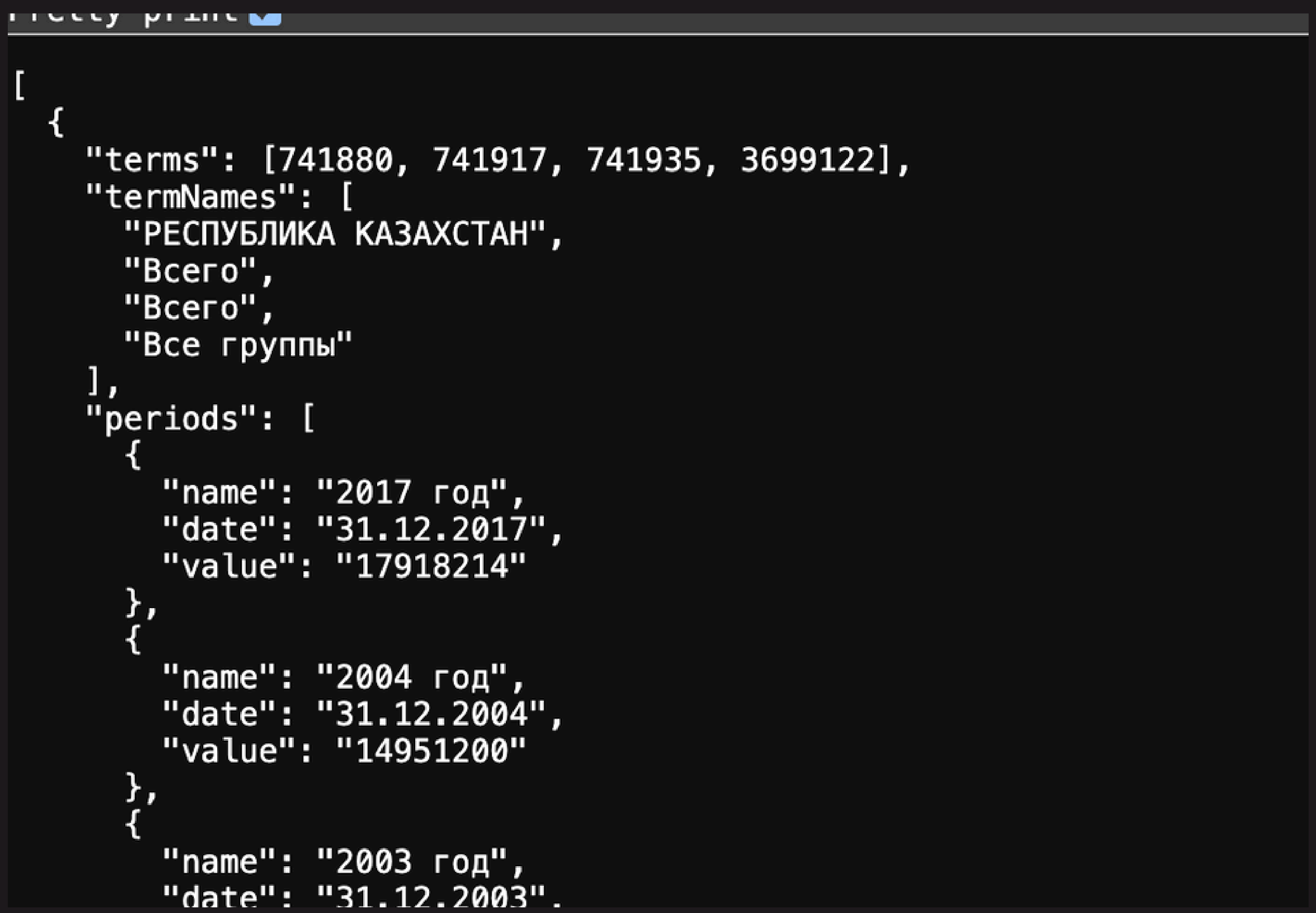
Некелер (қала/ауыл халқы)

Бұл жинақта 2000-2023 жылдардағы некелер саны (қала/ауыл халқы) туралы деректер бар

03.09.2024 0 128 Жарияланған

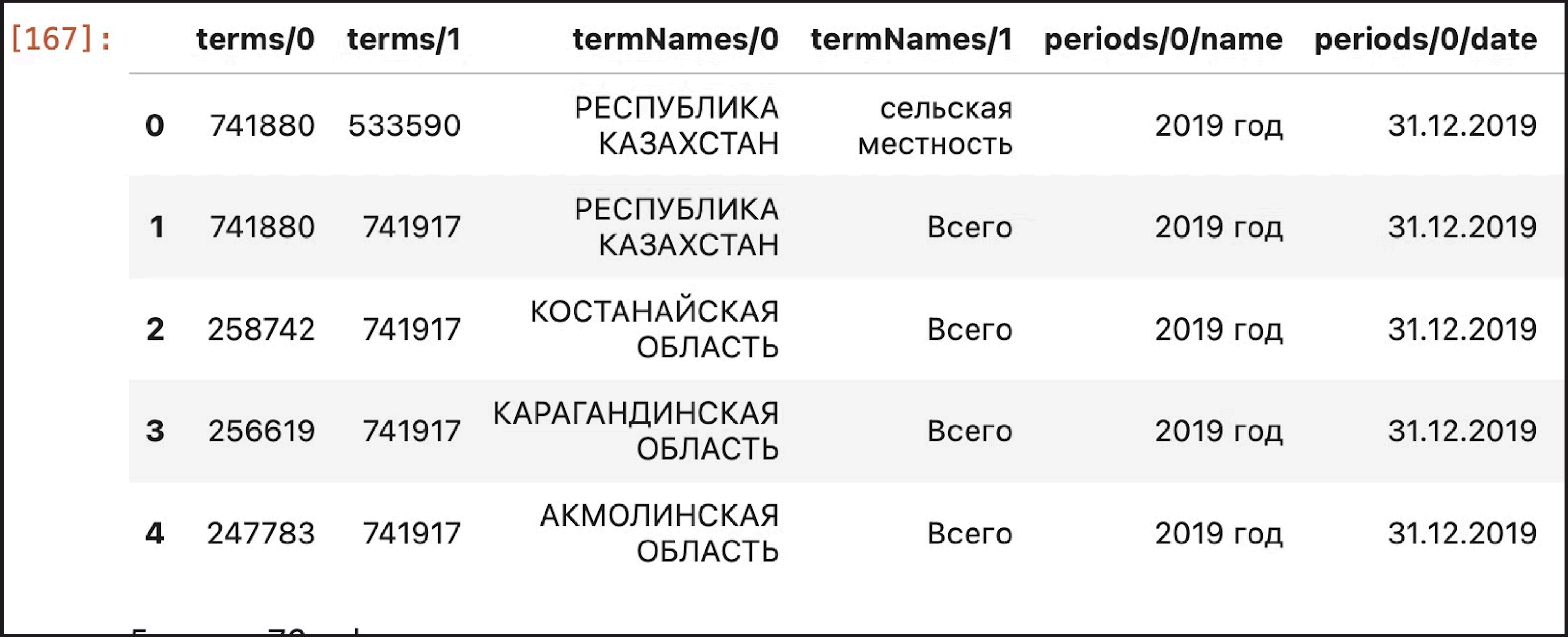
Data reformatting

convertcsv.com



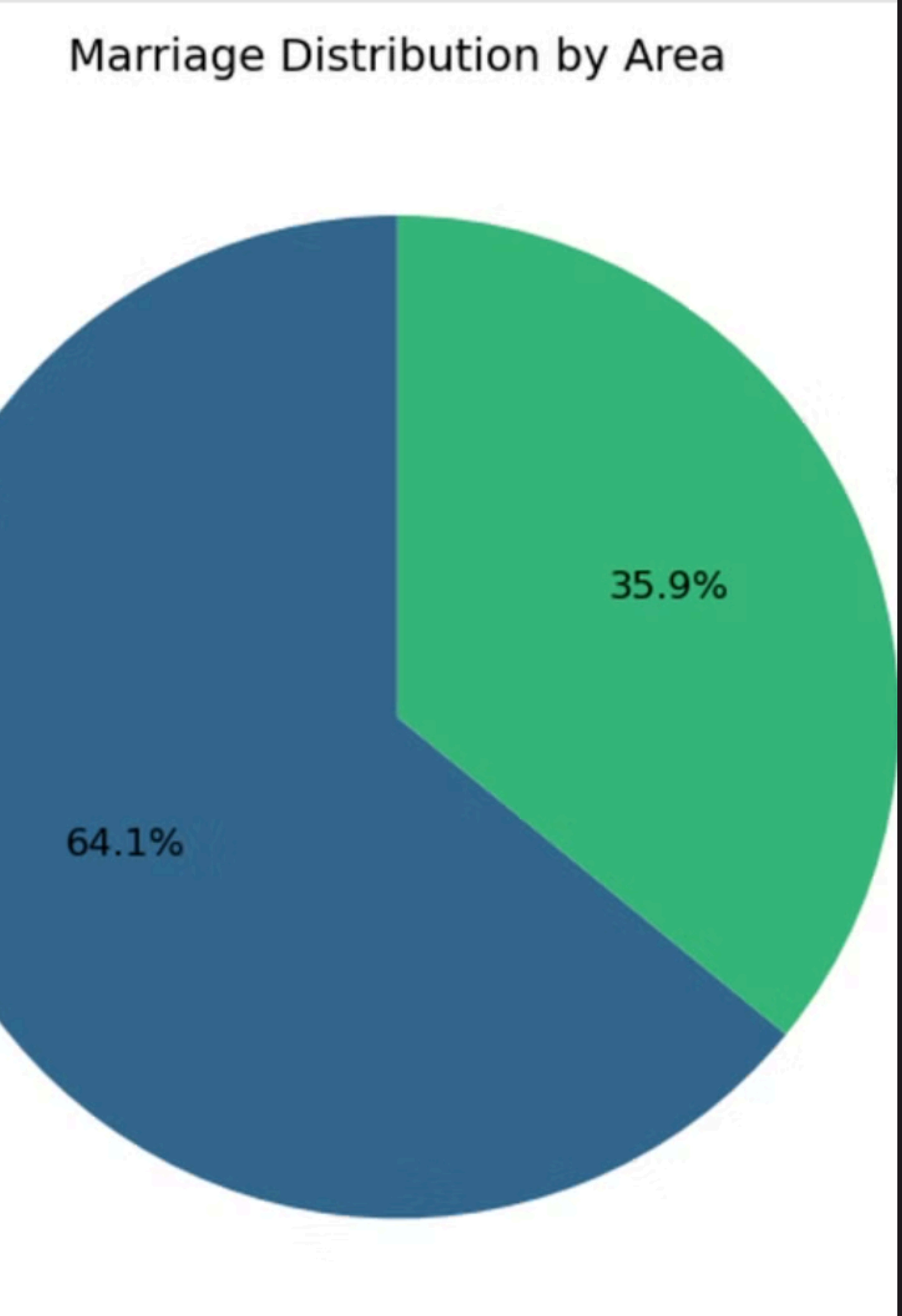
```
[
  {
    "terms": [741880, 741917, 741935, 3699122],
    "termNames": [
      "РЕСПУБЛИКА КАЗАХСТАН",
      "Всего",
      "Всего",
      "Все группы"
    ],
    "periods": [
      {
        "name": "2017 год",
        "date": "31.12.2017",
        "value": "17918214"
      },
      {
        "name": "2004 год",
        "date": "31.12.2004",
        "value": "14951200"
      },
      {
        "name": "2003 год",
        "date": "31.12.2003",
```

Data Preprocessing and Transformation

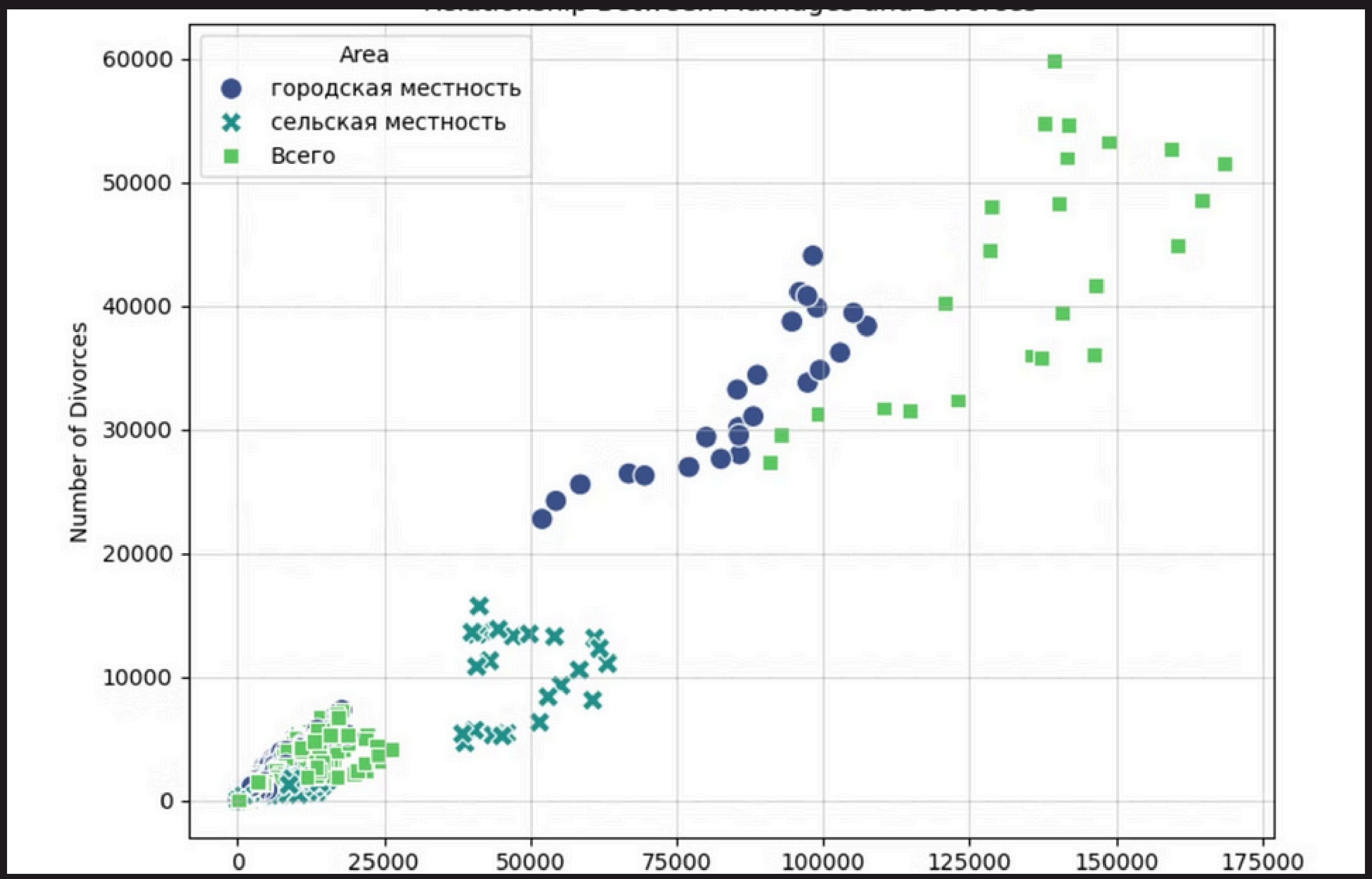


[167]:

	terms/0	terms/1	termNames/0	termNames/1	periods/0/name	periods/0/date
0	741880	533590	РЕСПУБЛИКА КАЗАХСТАН	сельская местность	2019 год	31.12.2019
1	741880	741917	РЕСПУБЛИКА КАЗАХСТАН	Всего	2019 год	31.12.2019
2	258742	741917	КОСТАНАЙСКАЯ ОБЛАСТЬ	Всего	2019 год	31.12.2019
3	256619	741917	КАРАГАНДИНСКАЯ ОБЛАСТЬ	Всего	2019 год	31.12.2019
4	247783	741917	АКМОЛИНСКАЯ ОБЛАСТЬ	Всего	2019 год	31.12.2019



Insights from data



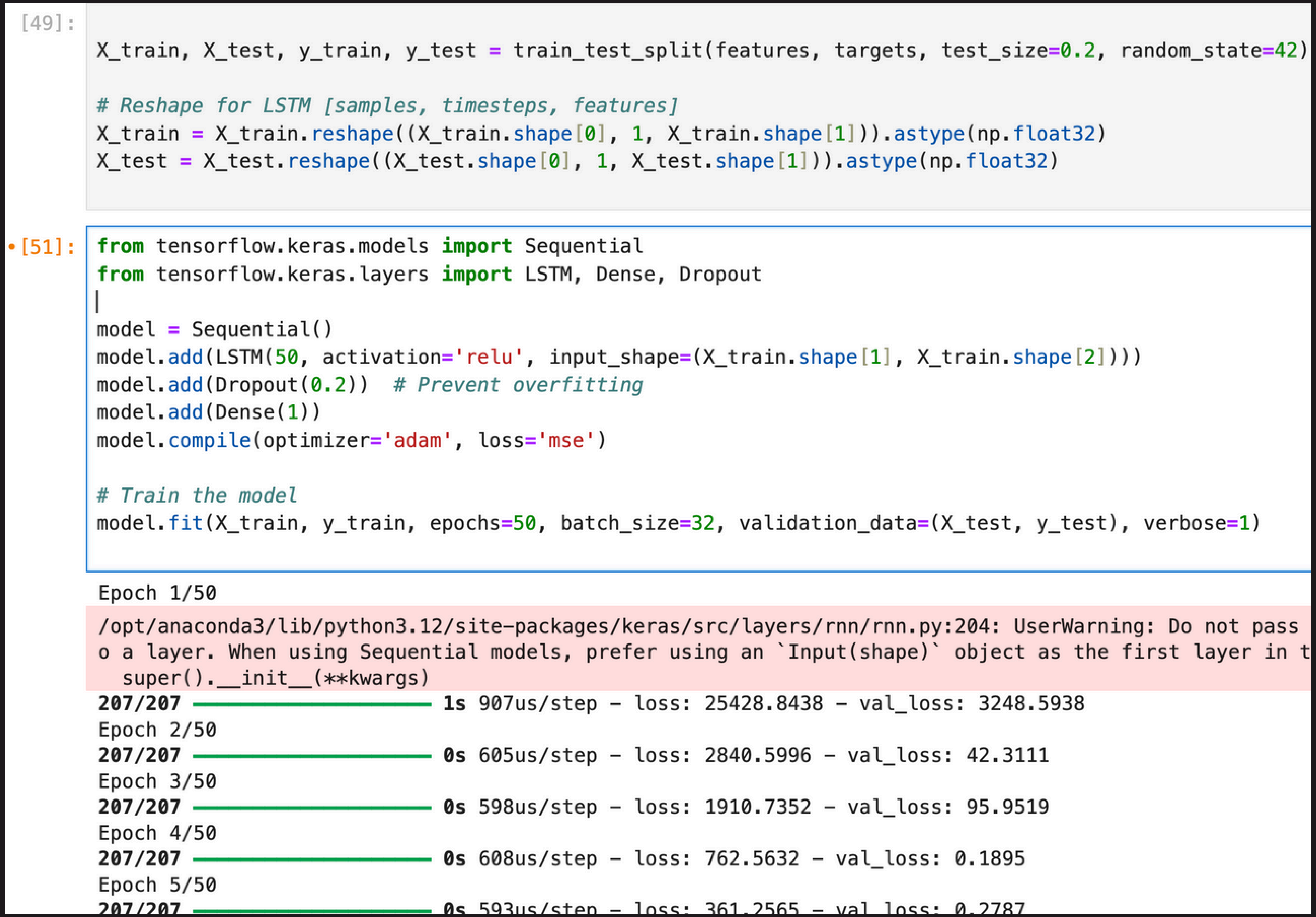
Highest Divorce Rate

Karaganda region has the highest divorce rate among the regions.

Highest Marriage Rate

Almaty region has the highest marriage rate among the regions.

Model Selection and Implementation

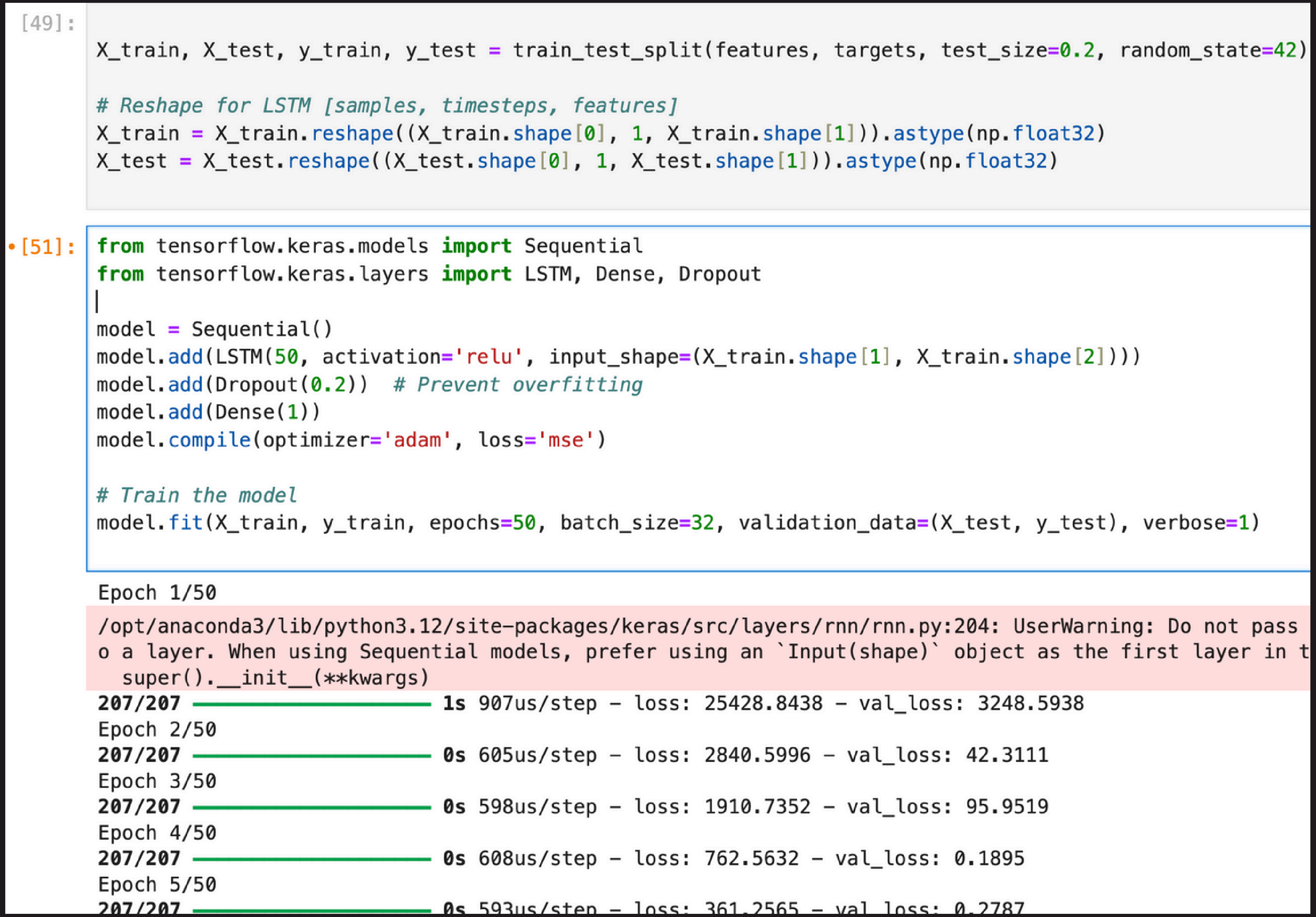


```
[49]:
X_train, X_test, y_train, y_test = train_test_split(features, targets, test_size=0.2, random_state=42)

# Reshape for LSTM [samples, timesteps, features]
X_train = X_train.reshape((X_train.shape[0], 1, X_train.shape[1])).astype(np.float32)
X_test = X_test.reshape((X_test.shape[0], 1, X_test.shape[1])).astype(np.float32)
```

```
•[51]:
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense, Dropout

model = Sequential()
model.add(LSTM(50, activation='relu', input_shape=(X_train.shape[1], X_train.shape[2])))
model.add(Dropout(0.2))  # Prevent overfitting
model.add(Dense(1))
model.compile(optimizer='adam', loss='mse')

# Train the model
model.fit(X_train, y_train, epochs=50, batch_size=32, validation_data=(X_test, y_test), verbose=1)
```

```
Epoch 1/50
/opt/anaconda3/lib/python3.12/site-packages/keras/src/layers/rnn/rnn.py:204: UserWarning: Do not pass
o a layer. When using Sequential models, prefer using an `Input(shape)` object as the first layer in t
  super().__init__(**kwargs)
207/207 ━━━━━━━━━━━━━━━━━━━━ 1s 907us/step - loss: 25428.8438 - val_loss: 3248.5938
Epoch 2/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 605us/step - loss: 2840.5996 - val_loss: 42.3111
Epoch 3/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 598us/step - loss: 1910.7352 - val_loss: 95.9519
Epoch 4/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 608us/step - loss: 762.5632 - val_loss: 0.1895
Epoch 5/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 593us/step - loss: 361.2565 - val_loss: 0.2787
```

Advanced Techniques and Analysis



```
[49]:
X_train, X_test, y_train, y_test = train_test_split(features, targets, test_size=0.2, random_state=42)

# Reshape for LSTM [samples, timesteps, features]
X_train = X_train.reshape((X_train.shape[0], 1, X_train.shape[1])).astype(np.float32)
X_test = X_test.reshape((X_test.shape[0], 1, X_test.shape[1])).astype(np.float32)
```

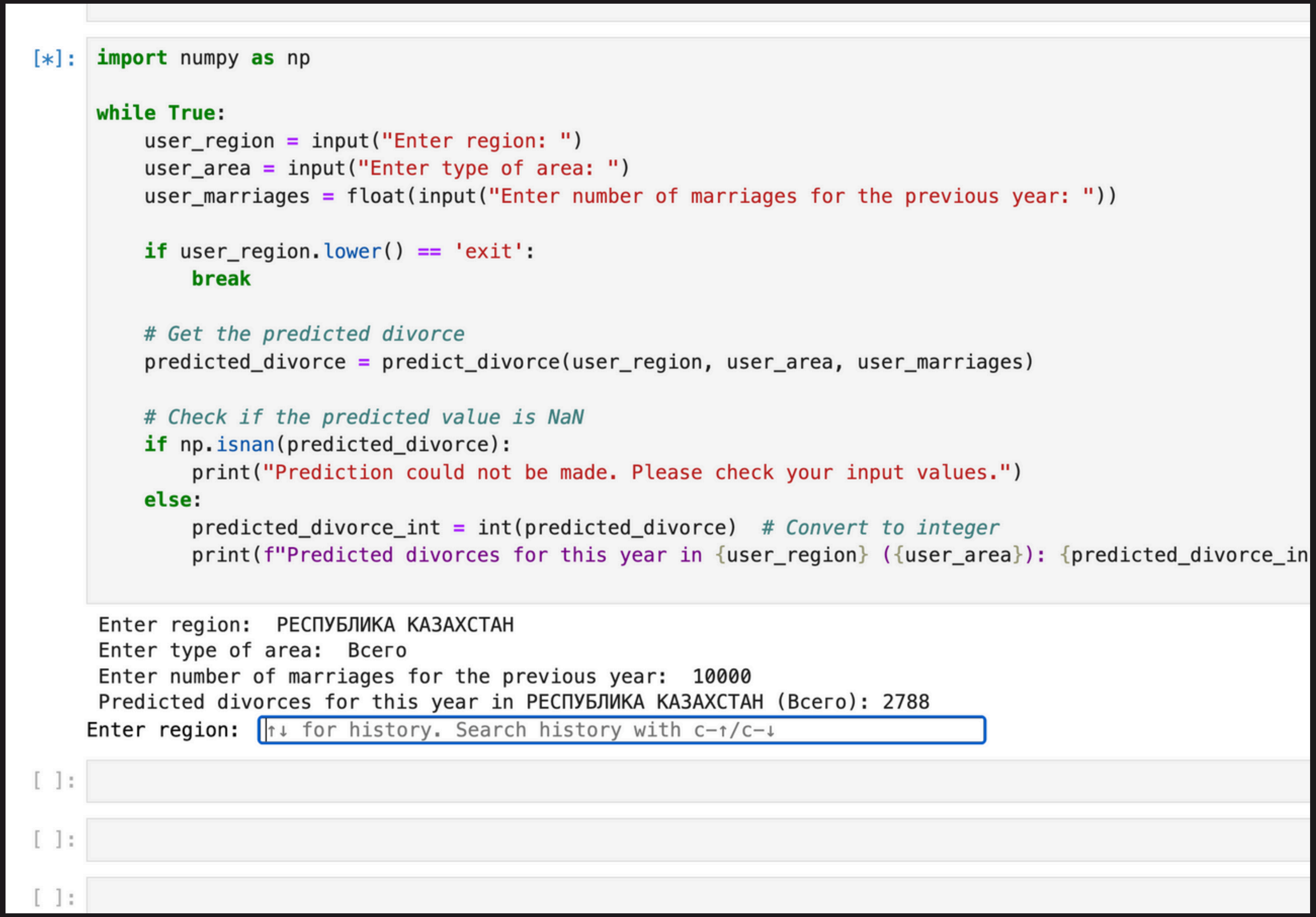
```
[51]:
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense, Dropout

model = Sequential()
model.add(LSTM(50, activation='relu', input_shape=(X_train.shape[1], X_train.shape[2])))
model.add(Dropout(0.2))  # Prevent overfitting
model.add(Dense(1))
model.compile(optimizer='adam', loss='mse')

# Train the model
model.fit(X_train, y_train, epochs=50, batch_size=32, validation_data=(X_test, y_test), verbose=1)
```

```
Epoch 1/50
/opt/anaconda3/lib/python3.12/site-packages/keras/src/layers/rnn/rnn.py:204: UserWarning: Do not pass
o a layer. When using Sequential models, prefer using an `Input(shape)` object as the first layer in t
  super().__init__(**kwargs)
207/207 ━━━━━━━━━━━━━━━━━━━━ 1s 907us/step - loss: 25428.8438 - val_loss: 3248.5938
Epoch 2/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 605us/step - loss: 2840.5996 - val_loss: 42.3111
Epoch 3/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 598us/step - loss: 1910.7352 - val_loss: 95.9519
Epoch 4/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 608us/step - loss: 762.5632 - val_loss: 0.1895
Epoch 5/50
207/207 ━━━━━━━━━━━━━━━━━━━━ 0s 593us/step - loss: 361.2565 - val_loss: 0.2787
```

MinMaxScaler()

Adaptive Moment Estimation (Adam)

Usage of the model



```
[*]: import numpy as np

     while True:
         user_region = input("Enter region: ")
         user_area = input("Enter type of area: ")
         user_marriages = float(input("Enter number of marriages for the previous year: "))

         if user_region.lower() == 'exit':
             break

         # Get the predicted divorce
         predicted_divorce = predict_divorce(user_region, user_area, user_marriages)

         # Check if the predicted value is NaN
         if np.isnan(predicted_divorce):
             print("Prediction could not be made. Please check your input values.")
         else:
             predicted_divorce_int = int(predicted_divorce)  # Convert to integer
             print(f"Predicted divorces for this year in {user_region} ({user_area}): {predicted_divorce_in
```

```
Enter region:  РЕСПУБЛИКА КАЗАХСТАН
Enter type of area:  Всего
Enter number of marriages for the previous year:  10000
Predicted divorces for this year in РЕСПУБЛИКА КАЗАХСТАН (Всего): 2788
Enter region: ↑↓ for history. Search history with c-↑/c-↓
```

```
[ ]:
[ ]:
[ ]:
```