

Descriptive Statistics EU vs EEU

Anton Yeshchenko, Kerwin Jorbina

24 April 2016

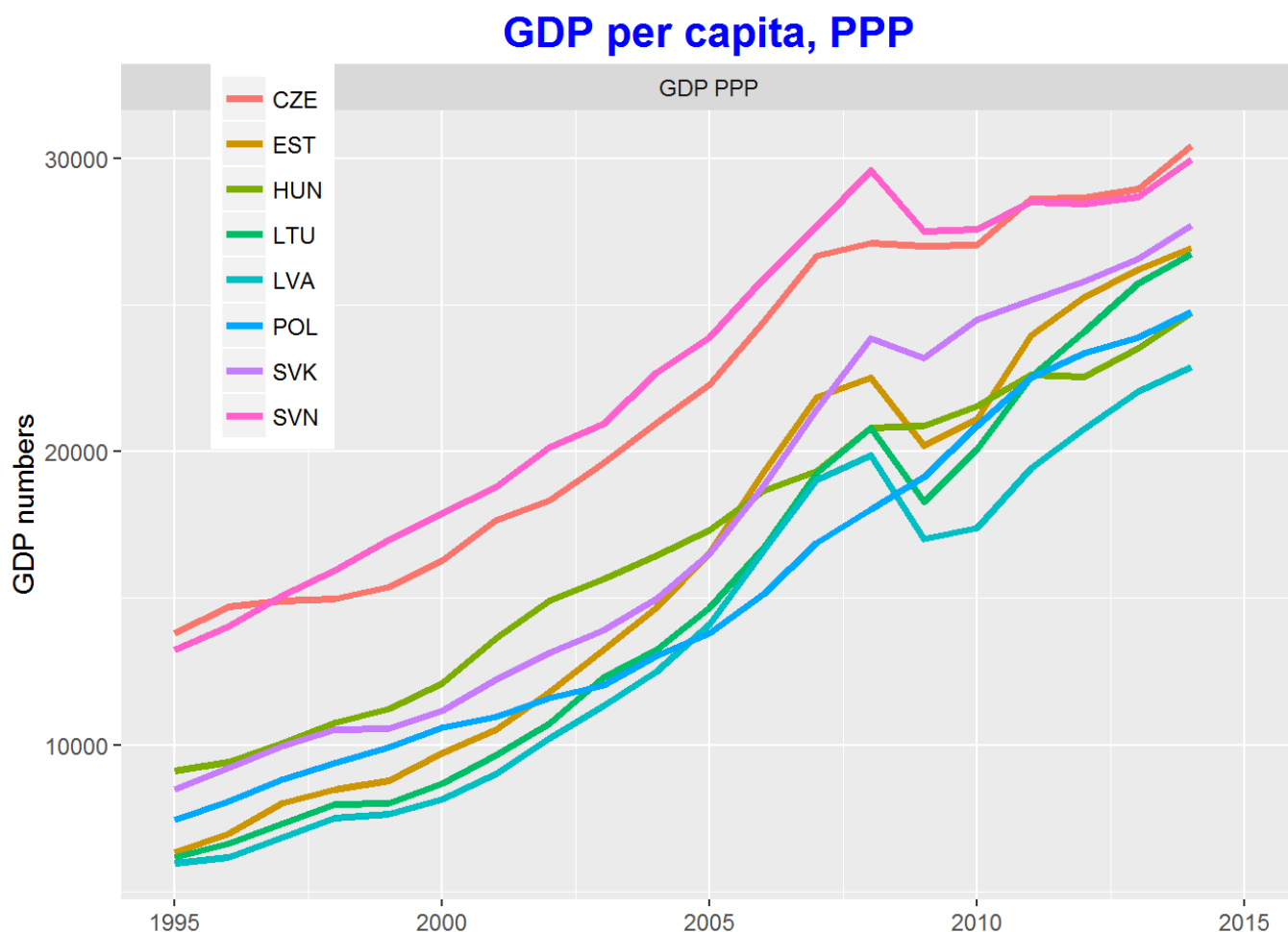
Descriptive statistics

We have a set that weights 500mb, and contains 1421 variables for every country. So we expect to spend quite some time on 'feature engineering', what is the best indicators of country's growth and betterness?

First, we start from looking into GDP, as the main indicator of economy growth.

EU countries (joined after 2004)

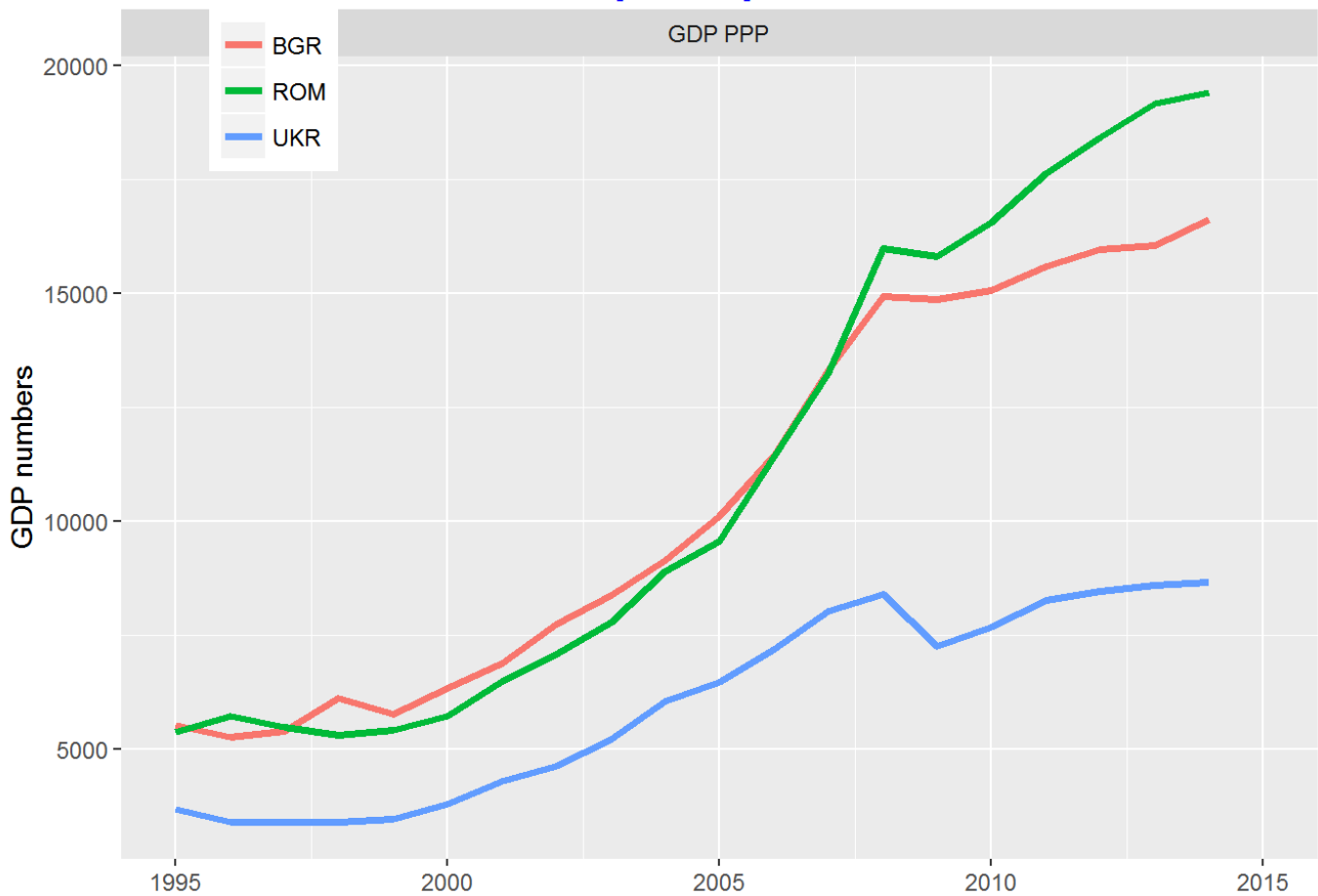
```
## Warning: Removed 8 rows containing missing values (geom_path).
```



EU countries (joined after 2007) + Ukraine

```
## Warning: Removed 3 rows containing missing values (geom_path).
```

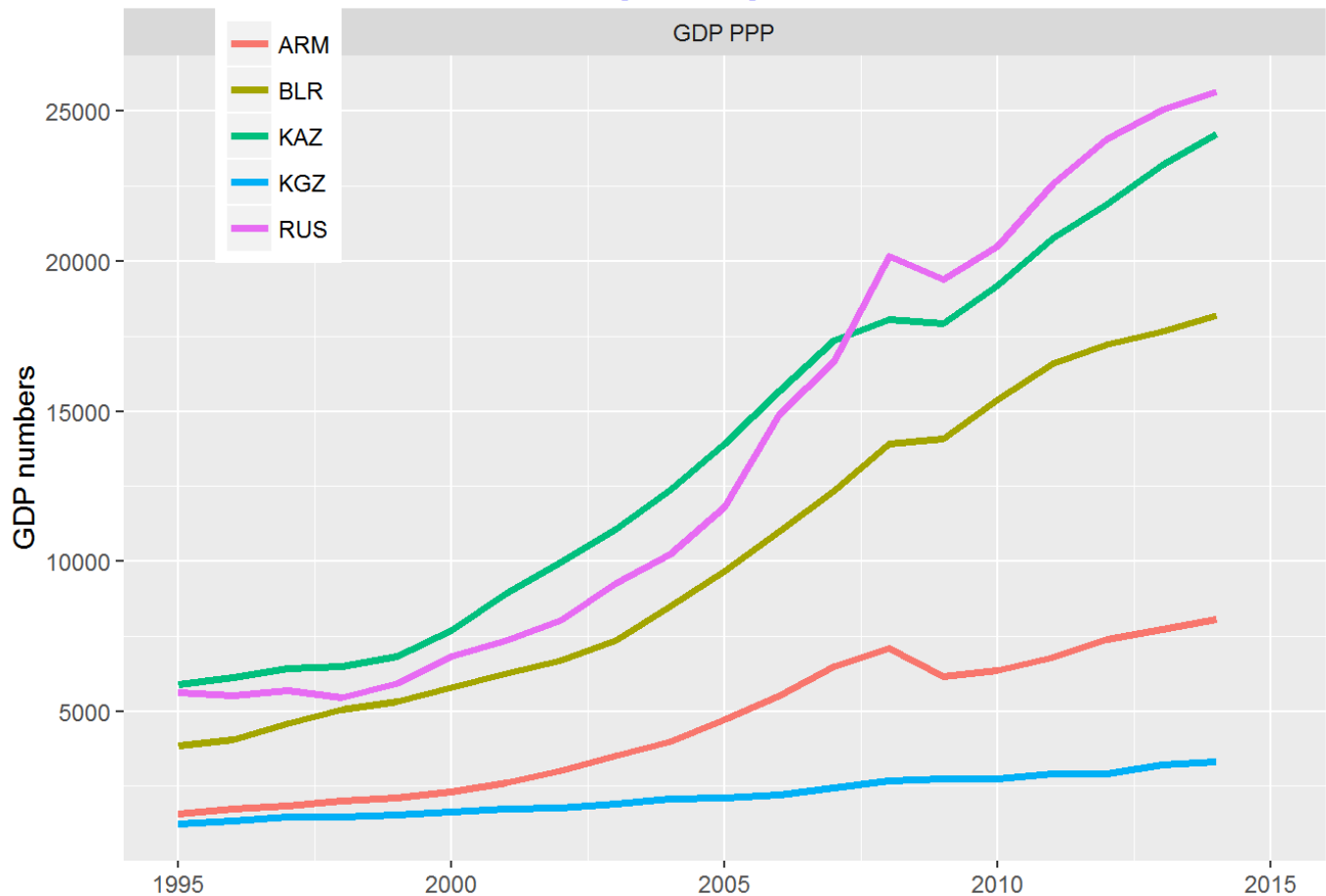
GDP per capita, PPP



EEU countries

```
## Warning: Removed 5 rows containing missing values (geom_path).
```

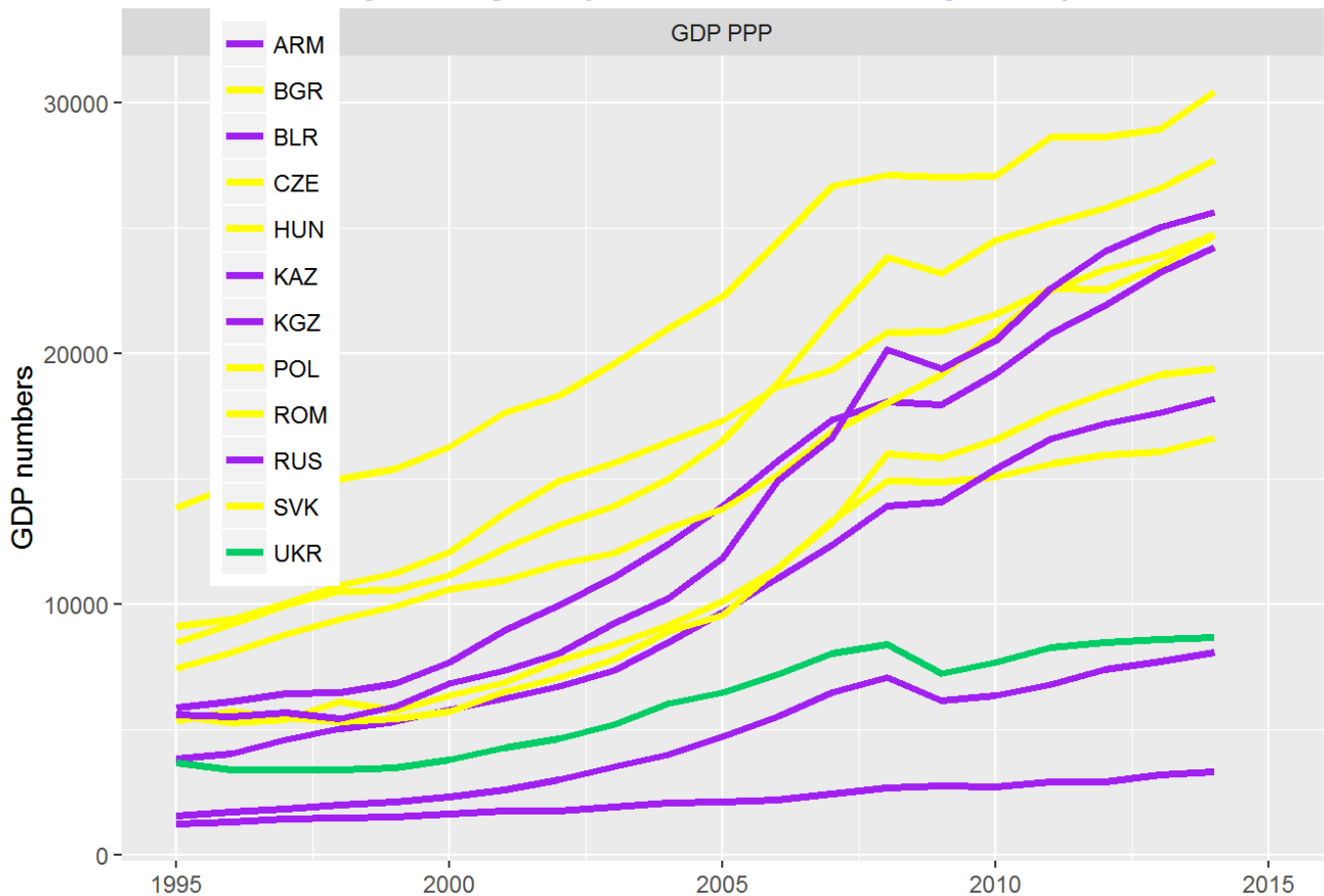
GDP per capita, PPP



Now all plots on the same scale, but colored by EU2004-EU2007-EEU

```
## Warning: Removed 12 rows containing missing values (geom_path).
```

GDP per capita (EU-EEU-Ukr compared), PPP



```
## [1] "Armenia"          "Belarus"          "Bulgaria"
## [4] "Czech Republic"   "Hungary"          "Kazakhstan"
## [7] "Kyrgyz Republic" "Poland"           "Romania"
## [10] "Russian Federation" "Slovak Republic" "Ukraine"
```

Now some unemployment)

```
## Warning in colnames(unemployment)[3] <- c("Total", "Female", "Male"):
## number of items to replace is not a multiple of replacement length
```

```
## Warning: Removed 12 rows containing missing values (geom_path).
```

UNEMPLOYMENT LEVEL

