

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное
учреждение высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Инженерная школа информационных технологий и робототехники

Направление – 09.04.01 Информатика и вычислительная техника

Отделение информационных технологий

«Импорт данных из международных баз данных»

Лабораторная работа №4 по дисциплине

«Проектирование программного обеспечения»

Выполнил студент гр. 8ВМ1И



Подпись

01.10.2021

Дата

Мунько А. С.

Проверил:

Подпись

Дата

Хамухин А. А.

Введение

Целью данной работы является научиться получать данных из различных публичных источников, а также визуализировать их.

Использованные библиотеки:

- dplyr;
- tidyr;
- ggplot2;
- eurostat;
- nasapower;
- soilDB.

Подготовка данных

Для загрузки данных используется следующий код, причем, некоторые параметры функций были изменены в новых версиях библиотек:

```
tables = c('tet00034', 'tet00033', 'tet00032', 'tet00031', 'tet00030',
'tet00029')

trades = lapply(tables, function(X) {
  get_eurostat(X) %>% label_eurostat()
}) %>%
  bind_rows() %>%
  select(-geo) %>%
  dplyr::filter(stringr::str_detect(indic_et, 'Exports in|Imports
in')) %>%
  pivot_wider(names_from = indic_et, values_from = values) %>%
  rename(
    export = `Exports in million of ECU/EURO`,
    import = `Imports in million of ECU/EURO`
  ) %>%
  mutate(partner = as.factor(partner))

daily_single_ag <- get_power(
  community = "AG",
  lonlat = c(60.59, 56.84),
  pars = c("RH2M", "T2M", "PRECTOTCORR"),
  dates = c("1995-04-01", "1995-04-30"),
  temporal_api = "DAILY"
)

interannual_sse <- get_power(
  community = "RE",
  lonlat = c(60.59, 56.84),
  pars = c("CLRSKY_SFC_SW_DWN", "ALLSKY_SFC_SW_DWN"),
  dates = 1995:2015,
  temporal_api = "MONTHLY"
)

soils = c('wilkes', 'chewacla', 'congaree')
```

```
series = fetchOSD(soils, extended = TRUE)

trades_total = trades %>%
  group_by(time) %>%
  summarise(export = sum(export), import = sum(import))
```

Визуализация данных

Ранее посчитанные данные – `trades_total` представляют собой данные торговли, сгруппированные по времени (Рис. 1).

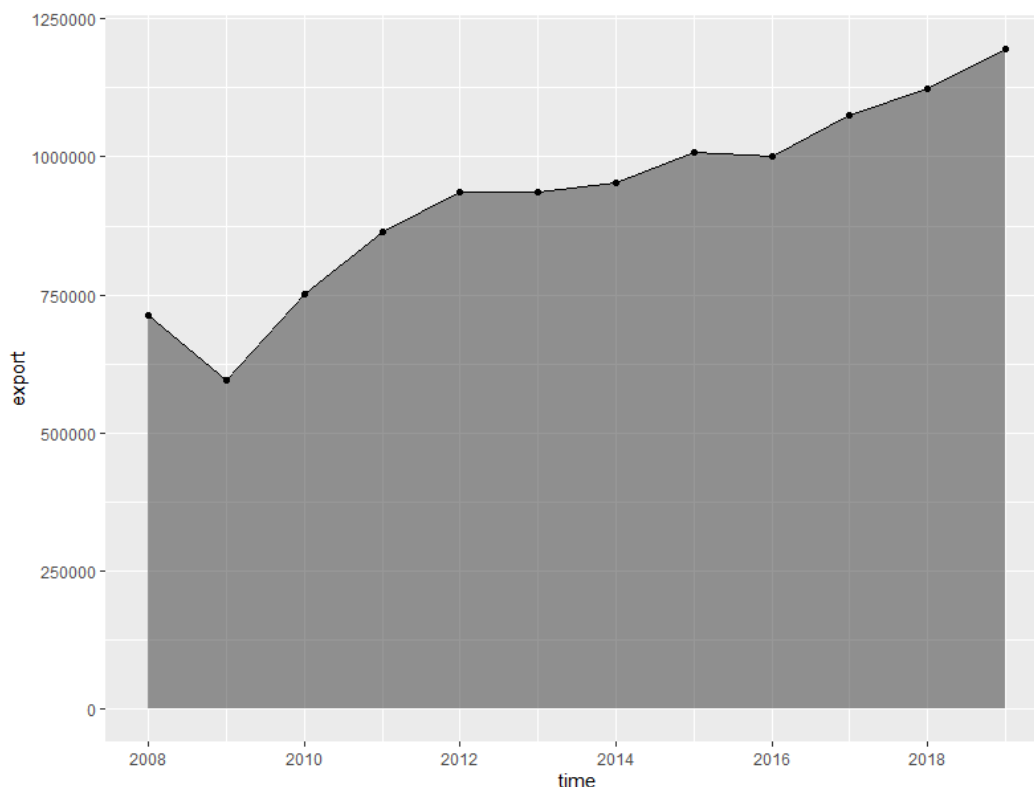


Рисунок 1 – Сгруппированные данные торговли.

Также, можно визуализировать данные за определенное время и определенную категорию торговли (Рис. 2).

Так как представленные данные отображают партнера экспорта, то можно использовать конкретную страну в выборке (Рис. 3):

```
trade_russia = trades %>% dplyr::filter(partner == 'Russia')

ggplot(trade_russia, mapping = aes(x = time, y = export, color = sitc06)) +
  geom_line(mapping = aes(group = sitc06)) +
  geom_point(mapping = aes(shape = sitc06)) +
  labs(x = "Год", y = "Экспорт", color = "Категория", shape = "Категория")
```

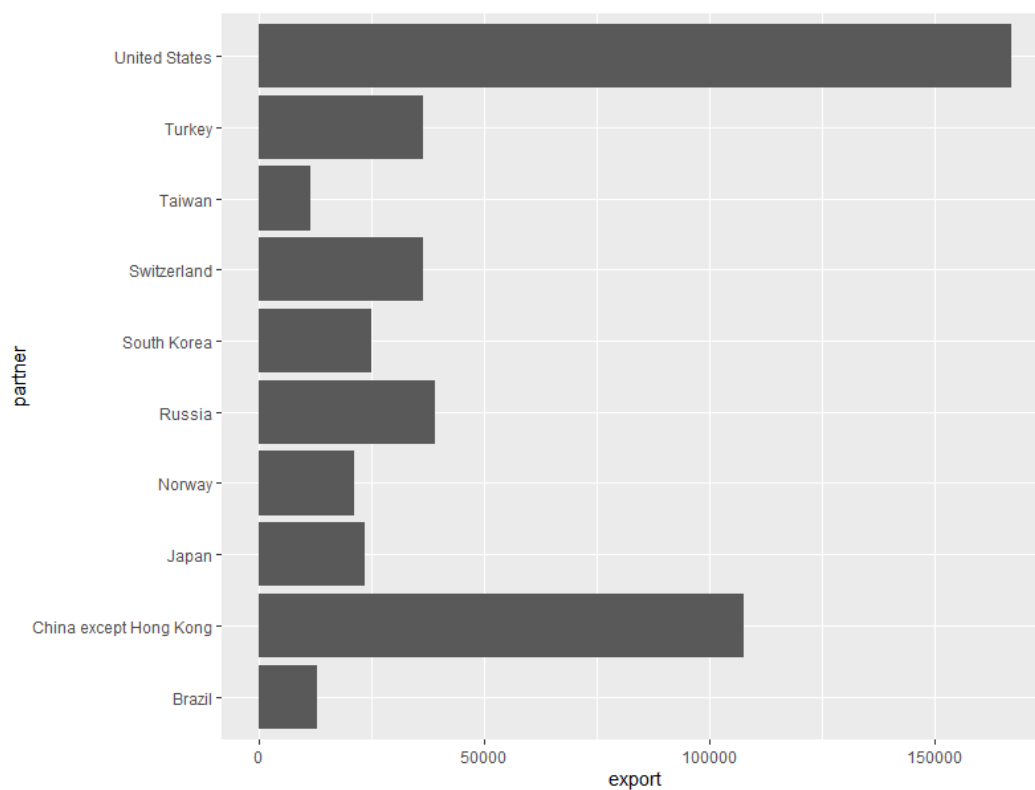


Рисунок 2 – Данные экспорта за январь 2017 года по категории Machinery and transport equipment.

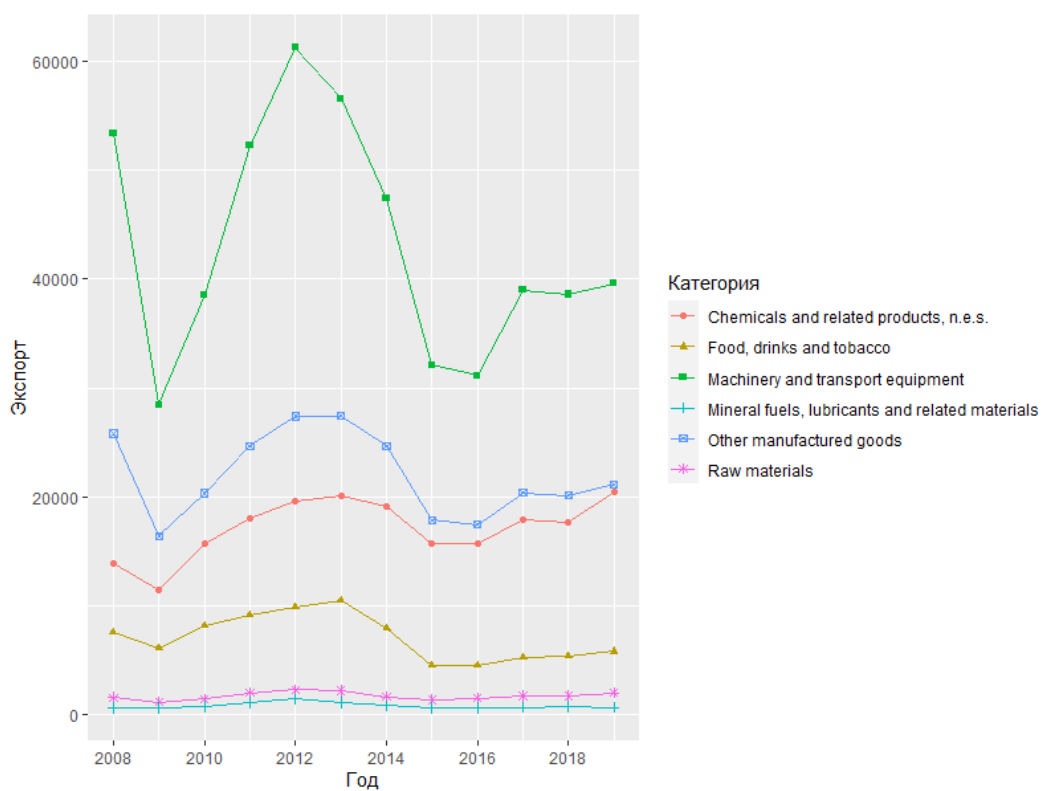


Рисунок 3 – Данные экспорта, в которых партнером фигурирует Россия.

Библиотека `ggplot2` позволяет использовать полярную систему координат (Рис. 4):

```
trades %>%
  dplyr::filter(sitc06 == 'Machinery and transport equipment', time ==
as.Date('2017-01-01')) %>%
  ggplot(mapping = aes(x = '', y = export, fill = partner), color =
'black', size = 0.2) +
  geom_col() +
  coord_polar(theta = 'y')
```

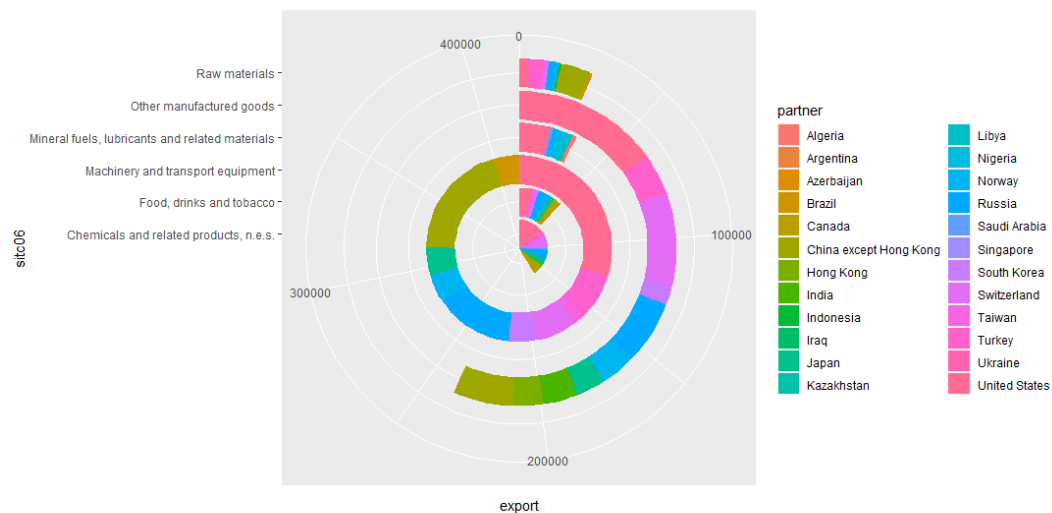


Рисунок 4 – Пример использования полярной системы координат.

Библиотека `ggplot2` позволяет создавать комплексные графики с использованием логарифмического масштабирования, цветной маркировкой и изменением размеров (Рис. 5):

```
trades_type = trades %>%
  group_by(sitc06, time) %>%
  summarise(export = sum(export), import = sum(import))

options(scipen = 999)

ggplot(trades_type) +
  geom_point(mapping = aes(x = export, y = import, color = sitc06,
size = time), alpha = 0.5) +
  scale_x_log10(breaks = seq(0, 500000, 100000)) +
  scale_y_log10(breaks = seq(0, 500000, 100000)) +
  labs(color = "Вид продукции", size = 'Год') +
  ggtitle('Соотношение импорта и экспорта в странах Евросоюза (млн
долл. США)', subtitle = 'Данные по ключевым партнерам') +
```

```
xlab('Экспорт') +  
ylab('Импорт')
```

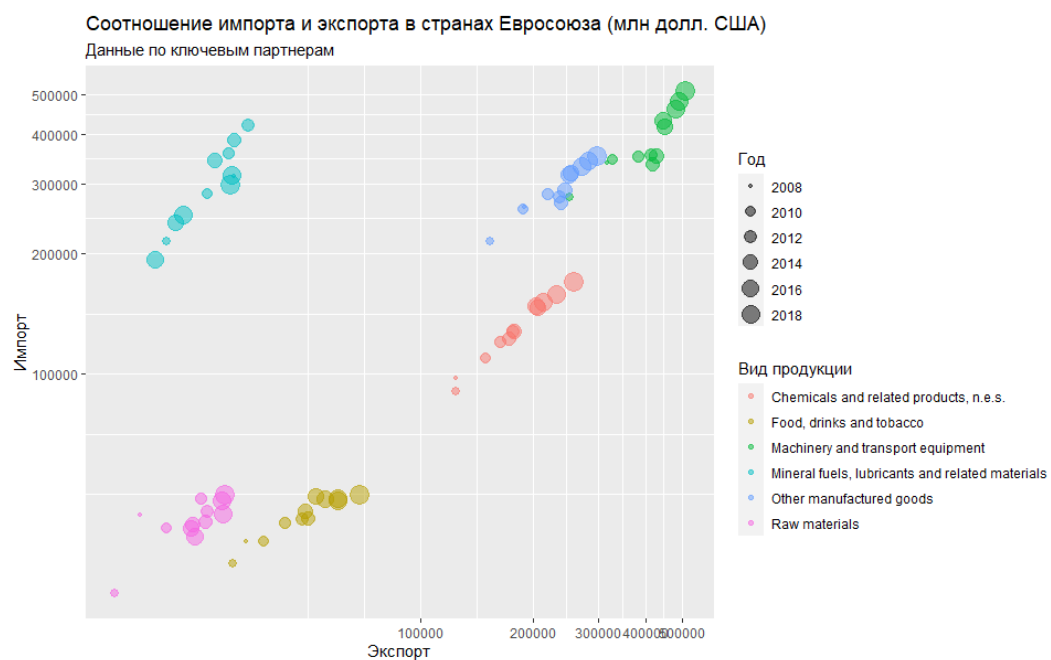


Рисунок 5 – Пример использования комплексных параметров при создании графика ggplot2.