Draft v2

1.0 Introduction

The United Kingdom's withdrawal from the European Union represents one of the most significant institutional and economic changes in recent decades. The Brexit referendum was held in June 2016, and the UK officially left the European Union (EU) on 31 January 2020, marking the end of nearly five decades of membership. EU has long been the UK's largest trading partner, Brexit introduced new barriers to trade, altered market access conditions, and reshaped the country's external economic relationships.

This project aims to analyse how the UK's trade with EU and non-EU partners has evolved before and after Brexit, with a particular focus on identifying structural shifts in trade intensity. The analysis is grounded in the gravity model of international trade, which provides data for explaining bilateral trade as a function of economic size, distance, and institutional linkages. By examining trade developments between 2010 and 2021*, we seek to explore whether Brexit has led to changes in trading patterns

1.1 Research question

*How has the United Kingdom's withdrawal from the European Union affected its trade flows with EU and non-EU partners?**

The objective is to document changes in the level and composition of UK goods trade before and after Brexit.

2.0 Data collection

The dataset draws on the CEPII Gravity Database and contains a targeted selection of bilateral variables between the United Kingdom and its trading partners. When uploading the gravity data set, we downloaded the Gravity documentation file which included information about the variables in the data set. Among all variables, we selected variables that present the main economic, geographic, and institutional factors that may have shaped trade patterns before and after Brexit.

2.1 Packages

```
# For data collection
library(here)

# For data manipulation
library(dplyr)
library(tidyverse)

# For making visualisations
library(ggplot2)
library(patchwork)      # helper package for side-by-side plots
library(scales)         # Scale functions for visualisations

# For report generation
library(knitr)

# Global settings
options(scipen = 999)     #makes R write large numbers without scientific notations
```

2.2 Creating the datasets

For this analysis we gathered data from the gravity dataset … we also gathered data from the World Development Indicators (WDI) from the World Bank.

```
# Read the gravity data
gravity <- readRDS(here("data", "Gravity_V202211.rds")) %>% # Gravity source file location
  select(
    year, iso3_o, iso3_d, eu_o, eu_d,
    country_exists_o, country_exists_d,
    distcap, pop_o, pop_d, gdp_o, gdp_d,
    tradeflow_comtrade_o, tradeflow_comtrade_d
  ) %>%
  filter(
    iso3_o == "GBR" | iso3_d == "GBR",        # only include observations with the UK
    country_exists_o == 1 | country_exists_d == 1,  # remove observations with non-existing c
    year %in% 2010:2020) %>%                  # focus on 2010-2020
  mutate(partner_eu = if_else(iso3_o == "GBR", eu_d, eu_o)) # EU-status for partnerland


# Read the WDI data
```

```
WDI <- read_csv("API_NY.GDP.PCAP.KD_DS2_en_csv_v2_242568.csv", col_names = TRUE, skip = 3) %
  select(!starts_with("...70"), # AI disclaimer (mystery column)
         -"Indicator Code") %>%
  pivot_longer(
    cols = matches("^(19|20)"), # AI disclaimer, regex?
    names_to = "year",
    values_to = "value"
  ) %>%
  pivot_wider( # AI disclaimer: explained the documentation
    names_from = "Indicator Name",
    values_from = "value"
  ) %>%
  rename("GDPpc" = "GDP per capita (constant 2015 US$)") %>%
  filter(year %in% 2010:2024) %>%
  mutate(year = as.numeric(year))
```

```
New names:
Rows: 266 Columns: 70
-- Column specification
-------------------------------------------------------- Delimiter: "," chr
(4): Country Name, Country Code, Indicator Name, Indicator Code dbl (65): 1960,
1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, ... lgl (1): ...70
i Use `spec()` to retrieve the full column specification for this data. i
Specify the column types or set `show_col_types = FALSE` to quiet this message.
* `` -> `...70`
```

Based on our main dataset, we also created smaller subsets of data … among them, seperate datasets containing UK exports and UK imports …

```
# All UK exports 2010-2020
uk_exports <- gravity %>%
  select(!starts_with("country_exists")) %>%
  filter(iso3_o == "GBR")

# All UK imports 2010-2020
uk_imports <- gravity %>%
  select(!starts_with("country_exists")) %>%
  filter(iso3_d == "GBR")

# Subsets by year
uk_exports_2010 <- uk_exports %>% filter(year == 2010)
uk_exports_2020 <- uk_exports %>% filter(year == 2020)
```

2.3 Variables

The trade flow variables (`tradeflow_comtrade_o`, `tradeflow_comtrade_d`) capture the value of exports and imports in current USD, forming the foundation for analysing how trade volumes evolved over time. To control for differences in market size and economic capacity, the dataset includes GDP (`gdp_o`, `gdp_d`) and population (`pop_o`, `pop_d`) for both the UK and its partners, providing context for the scale and potential intensity of bilateral trade. Geographic distance (`distcap`) serves as a proxy for trade costs, reflecting how physical separation can constrain trade through higher transport and transaction costs. The EU membership indicators (`eu_o`, `eu_d`) are central to the analysis, as they allow a distinction between EU and non-EU trading partners and thereby facilitate the identification of changes linked to the UK's withdrawal from the EU. Finally, the country existence variables ensure that only valid and contemporaneous country pairs are included each year. In total, these core variables were selected from the broader CEPII database to provide a balanced representation of economic scale, spatial frictions, and institutional context, enabling a systematic analysis of how the UK's trade flows may have been affected by Brexit.

3.0 Analysis

Countries were chosen because they are among the UK's trade partners, and they are all included in standard gravity datasets such as CEPII and UN Comtrade. They represent both nearby European economies and major global markets. Most have long-standing trade or historical ties with the UK, making them suitable for comparative analysis. Their inclusion ensures that it is possible to analyse bilateral trade composition and flows within a consistent gravity-model framework.

```r
trade_partners <- data.frame(
  Country = c(
    "Germany",
    "Netherlands",
    "France",
    "Italy",
    "Spain",
    "Belgium",
    "United States",
    "Switzerland",
    "Canada"
  ),
  `Key trade features` = c(
    "UK's largest EU trade partner, major machinery and automotive trade.",
    "Logistics hub (Rotterdam effect), re-exports dominate goods flows.",
    "Long-standing partner; key for manufacturing, transport, and defence industries.",
```

```
    "High-value goods, industrial machinery, and consumer products.",
    "Growing services and tourism linkages.",
    "Chemicals, pharmaceuticals, and transport equipment.",
    "The UK's largest single-country trading partner, with strong links in finance, technolog
    "Significant financial and pharmaceutical trade, with high-value exports and close invest
    "Expanding trade in energy, technology, and services, supported by the UK-Canada trade co
  )
)

kable(trade_partners, caption = "Overview: Key UK trade relationships by country included in
```

Table 1: Overview: Key UK trade relationships by country included in the analysis

Country	Key.trade.features
Germany	UK's largest EU trade partner, major machinery and automotive trade.
Netherlands	Logistics hub (Rotterdam effect), re-exports dominate goods flows.
France	Long-standing partner; key for manufacturing, transport, and defence industries.
Italy	High-value goods, industrial machinery, and consumer products.
Spain	Growing services and tourism linkages.
Belgium	Chemicals, pharmaceuticals, and transport equipment.
United States	The UK's largest single-country trading partner, with strong links in finance, technology, and services.
Switzerland	Significant financial and pharmaceutical trade, with high-value exports and close investment ties.
Canada	Expanding trade in energy, technology, and services, supported by the UK–Canada trade continuity agreement.

```
# Download GDP per capita (constant 2015 USD)
uk_gdp <- WDI %>%
  filter(`Country Code` == "GBR")



# Plot
ggplot(uk_gdp, aes(x = year, y = GDPpc )) +
  geom_line(color = "#003399", linewidth = 1.1) +
  geom_vline(xintercept = 2016, linetype = "dashed", color = "red") +
  geom_vline(xintercept = 2021, linetype = "dashed", color = "darkred") +
  labs(
    title = "UK GDP per capita (constant 2015 US$)",
    subtitle = "Dashed lines: 2016 Brexit referendum and 2021 Trade and Cooperation Agreemen
    x = "Year",
```

```
    y = "GDP per capita (USD, constant 2015)"
  ) +
  theme_minimal(base_size = 13)
```

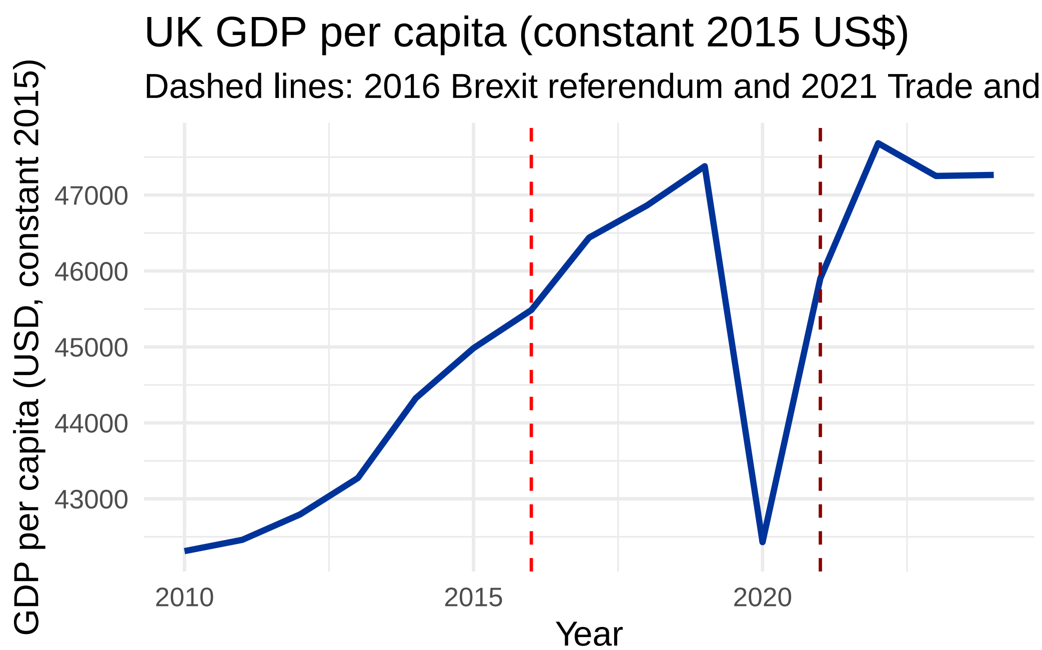


Figure 1: Figure: UK GDP per capita (constant 2015 US$), with 2016 referendum and 2021 Trade and Cooperation Agreement marked by dashed lines.

The first figure shows the evolution of UK real GDP per capita (constant 2015 USD) from 2000 to the present. The dashed vertical lines mark two key Brexit-related events: the 2016 referendum and the 2021 Trade and Cooperation Agreement (TCA). On its own, the UK series gives a sense of the country's long-run income trajectory, but it does **not** tell us whether Brexit had an effect, because we cannot see how the UK would have evolved in the absence of these shocks. For that, we need a comparison group.

The second figure therefore plots the cumulative change in real GDP per capita since 2010 for a set of advanced economies: Belgium, Canada, France, Germany, Italy, the Netherlands, Spain, Switzerland, and the UK. Using 2010 as a common baseline highlights how the UK's growth path compares with peers over the same period. The same dashed lines indicate the 2016 referendum and the 2021 TCA.

```
# Download GDP per capita (constant 2015 USD) for selected countries
dat <- WDI %>%
```

```r
  filter(`Country Code` %in% c("GBR", "DEU", "FRA", "ITA", "ESP", "NLD", "BEL", "CHN", "CAN")
  rename(country = `Country Code`)

# Calculate cumulative percent change since 2010
dat_growth <- dat %>%
  group_by(country) %>%
  mutate(
    base_2010 = GDPpc[year == 2010],
    pct_change = ((GDPpc / base_2010) - 1) * 100
  ) %>%
  ungroup()

# Plot with bold UK line
ggplot(dat_growth, aes(x = year, y = pct_change, group = country)) +
  geom_line(aes(color = country), alpha = 0.4, size = 1) +
  geom_line(
    data = filter(dat_growth, country == "United Kingdom"),
    aes(x = year, y = pct_change, color = country),
    size = 1.8
  ) +
  geom_vline(xintercept = 2016, linetype = "dashed", color = "red") +
  geom_vline(xintercept = 2021, linetype = "dashed", color = "darkred") +
  labs(
    title = "Cumulative change in real GDP per capita since 2010",
    subtitle = "Dashed lines: 2016 Brexit referendum and 2021 Trade and Cooperation Agreement
    x = "Year",
    y = "Percent change since 2010",
    color = "Country",
    caption = "Source: World Bank, World Development Indicators (accessed 2025-11-01)"
  ) +
  scale_y_continuous(labels = percent_format(scale = 1)) +
  theme_minimal(base_size = 13) +
  theme(
    plot.title = element_text(face = "bold"),
    plot.caption = element_text(hjust = 0.5, size = 10)
  )
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```

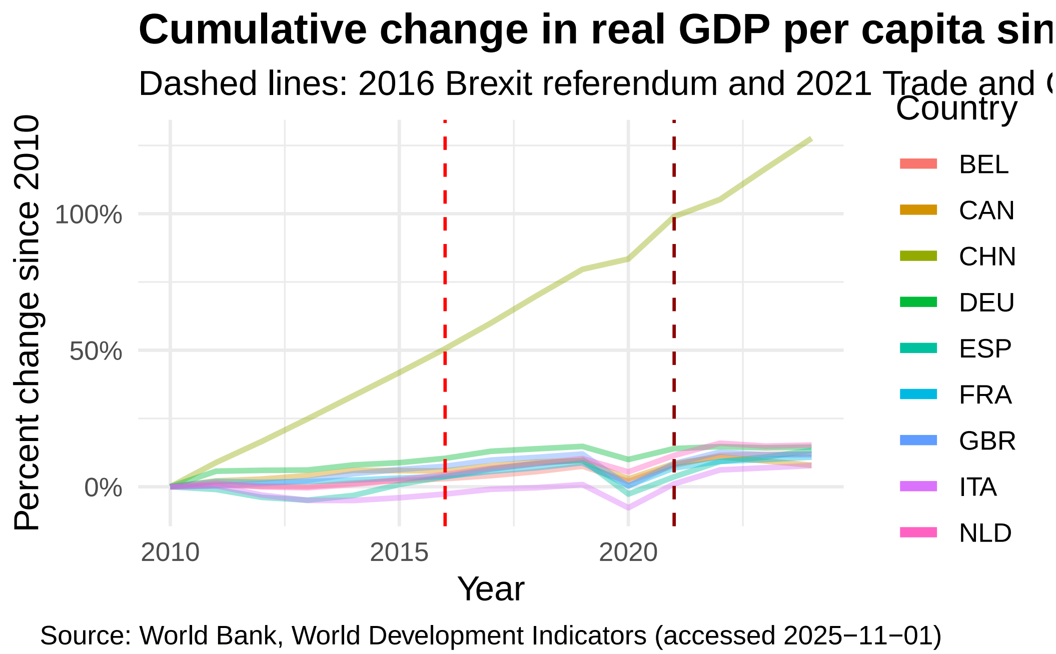


Figure 2: Figure: Cumulative change in real GDP per capita since 2010. Dashed lines mark the 2016 Brexit referendum and 2021 Trade and Cooperation Agreement.

To make the UK line appear bolder in the figure, the code first plots all countries using geom line () with lower opacity, alpha (0,4) and a moderate size. Then, added a second geom line layer that specifically filters the data for the United Kingdom (data = filter(dat_growth, country == "United Kingdom"), and plots it again with a thicker line, for a contrast effect.

From this comparative view, it becomes easier to evaluate relative performance. We can see whether the UK's cumulative growth accelerates or decelerates around the Brexit milestones compared with countries that were not directly exposed to Brexit-related trade shocks. While all countries experience a sharp dip around 2020 due to the COVID-19 pandemic.

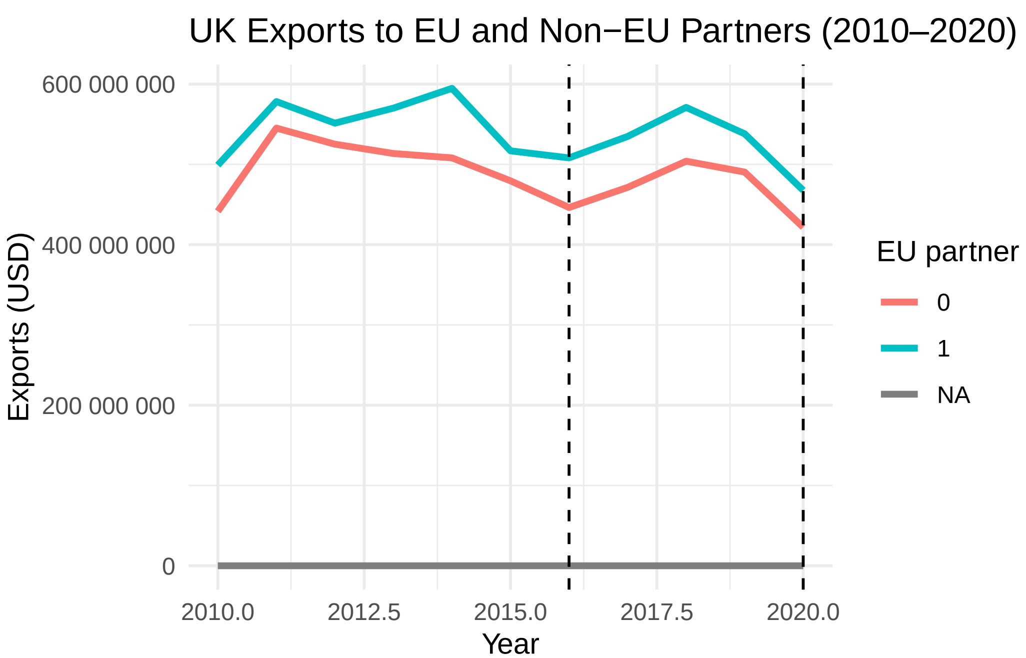
The row represents the single largest UK import flow in your dataset between 2010 and 2020. The variable iso3_o shows the origin country (DEU, meaning Germany) and iso3_d shows the destination country (GBR, the United Kingdom). The year variable indicates when this trade flow occurred. The import value itself is recorded in tradeflow_comtrade_d, which is the Comtrade-reported trade flow from Germany to the UK in that specific year. The variables pop_o and pop_d give the populations of the origin and destination countries, while gdp_o and gdp_d refer to their GDP levels. The variable distcap represents the distance between the countries' capital cities, and eu_o and eu_d indicate whether the origin and destination countries were members of the EU. All of these variables together describe the economic and geographic context of this specific trade flow, allowing you to analyse not only its value but also the conditions under which it occurred.

3.1 UK trade with EU and non-EU partners

```
uk_trade_eu <- gravity %>%
  group_by(year, partner_eu) %>%
  summarise(exports = sum(tradeflow_comtrade_o, na.rm = TRUE),
            imports = sum(tradeflow_comtrade_d, na.rm = TRUE),
            .groups = "drop")
```

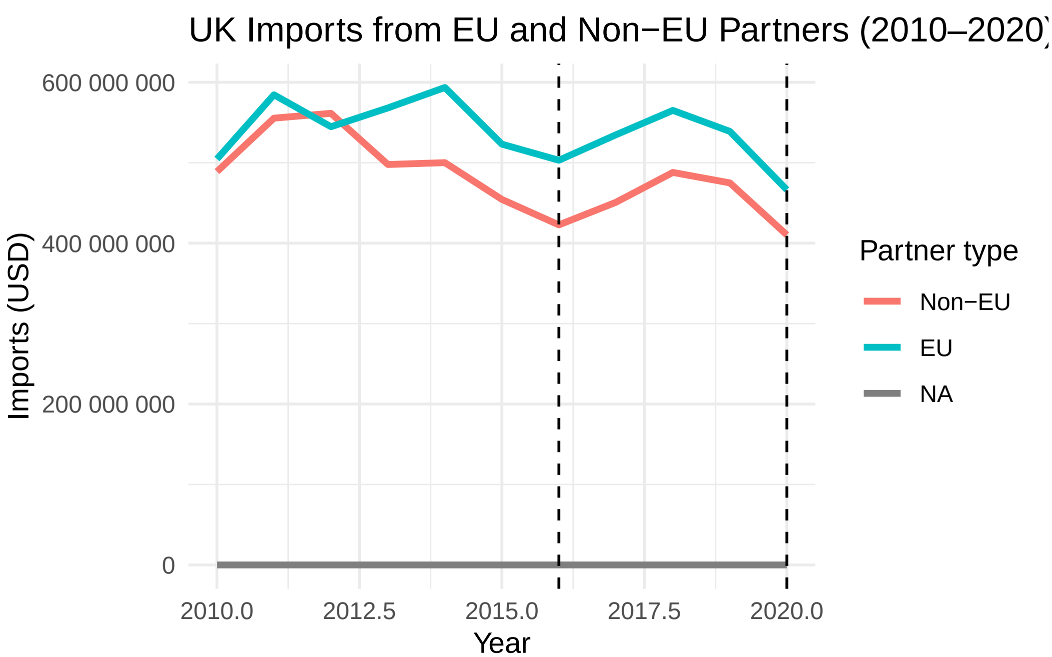
3.2 Plot: UK exports to EU vs non-EU

```
ggplot(uk_trade_eu, aes(x = year, y = exports, color = factor(partner_eu))) +
  geom_line(size = 1.2) +
  geom_vline(xintercept = 2016, linetype = "dashed") +
  geom_vline(xintercept = 2020, linetype = "dashed") +
  labs(title = "UK Exports to EU and Non-EU Partners (2010-2020)",
       x = "Year",
       y = "Exports (USD)",
       color = "EU partner") +
  scale_y_continuous(labels = scales::label_number()) +
  theme_minimal()
```



3.3 Plot: UK imports from EU vs non-EU

```
ggplot(uk_trade_eu, aes(x = year, y = imports, color = factor(partner_eu, labels = c("Non-EU"
)) +
  geom_line(size = 1.2) +
  geom_vline(xintercept = 2016, linetype = "dashed") +
  geom_vline(xintercept = 2020, linetype = "dashed") +
  labs(
    title = "UK Imports from EU and Non-EU Partners (2010-2020)",
    x = "Year",
    y = "Imports (USD)",
    color = "Partner type"
  ) +
  scale_y_continuous(labels = scales::label_number()) +
  theme_minimal()
```



The figure compares the UK's imports from EU and non-EU partners between 2010 and 2020. Imports from EU partners are consistently higher than imports from non-EU countries throughout the period, reflecting the UK's strong integration with the EU single market before Brexit. Both series show a dip around 2020 due to the COVID shock. There is no immediate visible drop following the 2016 referendum, but the flat or slightly declining trend after 2016 suggests that import growth from EU partners slowed relative to the pre-Brexit period.

3.4 EU share of UK trade

```
eu_share <- uk_trade_eu %>%
  group_by(year) %>%
  mutate(
    total_exports = sum(exports),
    total_imports = sum(imports),
    eu_share_exports = exports[partner_eu == 1] / total_exports,
    eu_share_imports = imports[partner_eu == 1] / total_imports
    ) %>%
  distinct(year, eu_share_exports, eu_share_imports)
```

```
Warning: There were 22 warnings in `mutate()`.
The first warning was:
i In argument: `eu_share_exports = exports[partner_eu == 1]/total_exports`.
i In group 1: `year = 2010`.
Caused by warning in `exports[partner_eu == 1] / total_exports`:
! longer object length is not a multiple of shorter object length
i Run `dplyr::last_dplyr_warnings()` to see the 21 remaining warnings.
```

Plot: EU Shares

```
ggplot(eu_share, aes(x = year)) +
  geom_line(aes(y = eu_share_exports, color = "Exports"), size = 1.2) +
  geom_line(aes(y = eu_share_imports, color = "Imports"), size = 1.2) +
  geom_vline(xintercept = 2016, linetype = "dashed") +
  geom_vline(xintercept = 2020, linetype = "dashed") +
  scale_y_continuous(labels = percent_format()) +
  labs(title = "EU Share of UK Exports and Imports",
       x = "Year",
       y = "EU share",
       color = "Flow"
       ) +
  theme_minimal()
```

```
Warning: Removed 1 row containing missing values or values outside the scale range
(`geom_line()`).
Removed 1 row containing missing values or values outside the scale range
(`geom_line()`).
```

The figure shows that the EU's share of UK exports and imports gradually declines after the 2016 referendum, with a sharper drop around 2020. This indicates that the UK becomes less dependent on EU trade over time, especially following the formal exit from the EU.

3.5 Largest UK export and import flows (2010–2020)

```
largest_export <- uk_exports %>%
  filter(!is.na(tradeflow_comtrade_o)) %>%
  slice_max(tradeflow_comtrade_o, n = 1) %>%
  select(year, iso3_d, tradeflow_comtrade_o)

largest_import <- uk_imports %>%
  filter(!is.na(tradeflow_comtrade_d)) %>%
  slice_max(tradeflow_comtrade_d, n = 1) %>%
  select(year, iso3_o, tradeflow_comtrade_d)

kable(largest_export, caption = "Largest UK export flow (2010-2020)")
```

Table 2: Largest UK export flow (2010–2020)

year	iso3_d	tradeflow_comtrade_o
2019	USA	71233629

```
kable(largest_import, caption = "Largest UK import flow (2010-2020)")
```

Table 3: Largest UK import flow (2010–2020)

year	iso3_o	tradeflow_comtrade_d
2014	DEU	99822005

```
summary(uk_exports$tradeflow_comtrade_o)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's 
       2    12201    70774  1970472   659921 71233629      331 
```

The first table shows which country the UK exported the most to in any single year between 2010 and 2020, and the size of that export flow. The second table shows the country from which the UK imported the most in any single year over the same period. These peak flows highlight the UK's strongest bilateral trade relationships, typically reflecting major EU partners such as Germany or the Netherlands for imports, and large markets like the United States or Germany for exports

3.6 Regression analysis

```
# Prepare data
regdata <- gravity %>%
  mutate(
    ln_exports = log(tradeflow_comtrade_o + 1),
    post2016 = if_else(year >= 2016, 1, 0)
  )

# Run model
model <- lm(ln_exports ~ partner_eu * post2016, data = regdata)

# Show regression results in the knitted report
summary(model)
```

```
Call:
lm(formula = ln_exports ~ partner_eu * post2016, data = regdata)

Residuals:
     Min       1Q   Median       3Q      Max 
-10.7664  -1.6649   0.0407   2.1915   7.3102 

Coefficients:
                    Estimate Std. Error t value            Pr(>|t|)    
(Intercept)         10.82745    0.06879 157.391 <0.0000000000000002 ***
partner_eu           4.09253    0.18570  22.038 <0.0000000000000002 ***
post2016            -0.03714    0.10313  -0.360               0.719    
partner_eu:post2016 -0.02956    0.27445  -0.108               0.914    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.071 on 4174 degrees of freedom
  (1355 observations deleted due to missingness)
Multiple R-squared:  0.1757,	Adjusted R-squared:  0.1751 
F-statistic: 296.6 on 3 and 4174 DF,  p-value: < 0.00000000000000022
```

4.0 Bibliography

Conte, M., P. Cotterlaz and T. Mayer (2022), "*The CEPII Gravity database*". CEPII Working Paper N°2022-05, July 2022.