Exam

Spatial and Temporal Climate Stress and Economic Indicators in Serbia

Jelena Cesnokova, 0001130702

November 1, 2024

Contents

Section A: Research Proposal

Analyzing Spatial and Temporal Climate Stress and Economic Indicators in Serbia (2015-2020)

Objective

This research aims to investigate the spatial and temporal variations in climate stress indicators and economic performance across Serbian NUTS 3 regions between 2015 and 2020. Specifically, the study focuses on the Standardized Precipitation-Evapotranspiration Index (SPEI) and the Water Stress Index, examining their correlation with economic indicators such as Agricultural Gross Value Added (GVA). This analysis will offer insights into how climate variability may impact economic outcomes across Serbian regions.

Research Questions

1. How have the SPEI and Water Stress Index evolved across Serbian NUTS 3 regions from 2015 to 2020?
2. What are the regional growth rates of these indices, and how do they spatially vary?
3. How do climate indicators like SPEI and water stress correlate with Agricultural GVA, and are there spatial spillover effects influencing neighboring regions?

Data and Methods

- Data Sources:
  - *Climate Data*: SPEI data from NASA Earthdata, processed into annual means for each year between 2015 and 2020. Water Stress Index data generated based on baseline conditions and simulated with random multipliers.
  - *Economic Data*: Agricultural GVA for Serbian NUTS 3 regions.
  - *Geospatial Data*: NUTS 3 shapefiles for Serbia, delineating regional boundaries.
- Data Preparation:
  - *Raster Manipulation*: Crop, resample, and align climate data with Serbian NUTS 3 boundaries.
  - *Time Series Aggregation*: Compute yearly averages for SPEI and water stress indices to observe temporal patterns.
  - *Spatial Data Management*: Integrate SPEI, water stress, and GVA data with NUTS 3 geometries for spatial analysis.

Analysis

1. Growth Rate Calculations:
   Calculate the growth rates for SPEI, water stress, and Agricultural GVA from 2015 to 2020 to assess spatial variations in climate stress and economic performance.

2. Mapping and Visualization:
   Use visualization tools (e.g., `ggplot2`, `sf`, `tmap` in R) to create spatial maps illustrating growth rates of SPEI, water stress, and Agricultural GVA. Temporal trends will also be visualized through time series graphs for each indicator across regions.

3. Spatial Econometric Analysis:

   - *Spatial Weighting*: Construct spatial weights matrices to define neighboring relationships between NUTS 3 regions.

Expected Outcomes

- Visual Insights:
  Spatial maps showing the growth rates of climate indices and Agricultural GVA will reveal regions with significant climate stress or economic variability. Time series plots will provide insights into temporal patterns in SPEI, water stress, and GVA across regions.

- Spatial Dependencies:
  Results from spatial econometric analysis will shed light on spatial dependencies, identifying whether climate stress in one region affects neighboring regions' economic performance. This will be particularly relevant for policy planning in agriculture and water resource management.

Significance

This research will contribute to a better understanding of the correlation between climate stress indicators and regional economic performance in Serbia. Insights from the study can guide targeted interventions to mitigate adverse climate impacts, particularly in vulnerable regions, enhancing resilience in agriculture and sustainable resource management.

Load Data

```r
library(ggplot2)
library(ncdf4)
library(terra)
library(spdep)
library(tmap)
library(raster)
library(sf)
library(spData)
library(tidyverse)
```

```r
shp2 <- st_read("C:/Users/jelce/Documents/Study/GIS/Exam/NUTS_RG_01M_2021_4326.shp")
```

```
## Reading layer `NUTS_RG_01M_2021_4326' from data source
##   `C:\Users\jelce\Documents\Study\GIS\Exam\NUTS_RG_01M_2021_4326.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 2010 features and 9 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: -63.15119 ymin: -21.38885 xmax: 55.83578 ymax: 80.83402
## Geodetic CRS:  WGS 84
```

Task 1

To start, we load the shapefile containing country boundaries and filter it to isolate Serbia. We then repeat this process to select the European NUTS 3 regions. In line with the exercise requirements, we proceed by creating a bounding box around the area of interest

```r
shp = world %>% filter(name_long == "Serbia")
serbia_reg <- st_read("C:/Users/jelce/Documents/Study/GIS/Exam/NUTS_RG_01M_2021_4326_LEVL_3.shp")
```

```
## Reading layer `NUTS_RG_01M_2021_4326_LEVL_3' from data source
##   `C:\Users\jelce\Documents\Study\GIS\Exam\NUTS_RG_01M_2021_4326_LEVL_3.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 1514 features and 9 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: -63.15119 ymin: -21.38885 xmax: 55.83578 ymax: 80.83402
## Geodetic CRS:  WGS 84
```

```r
shp3 = serbia_reg %>% filter(CNTR_CODE == "RS")
box <- st_bbox(shp)
box <- box+c(-.5,-.5,.5,.5)
```

Task 2

Then we load a raster object with SPEI data andr crop it.

```r
r <- rast("spei01.nc")
r
```

```
## class       : SpatRaster
## dimensions  : 360, 720, 1440  (nrow, ncol, nlyr)
## resolution  : 0.5, 0.5  (x, y)
## extent      : -180, 180, -90, 90  (xmin, xmax, ymin, ymax)
## coord. ref. : lon/lat WGS 84 (CRS84) (OGC:CRS84)
## source      : spei01.nc
## varname     : spei (Standardized Precipitation-Evapotranspiration Index)
## names       :   spei_1,   spei_2,   spei_3,   spei_4,   spei_5,   spei_6, ...
## unit        : z-values, z-values, z-values, z-values, z-values, z-values, ...
## time (days) : 1901-01-16 to 2020-12-16
```

```
inMemory(r)
```

```
## [1] FALSE
```

We examined the structure of the raster and identified that layers 1369 to 1440 correspond to the last six years, as the data is organized by months. To facilitate easier analysis, we aggregated the data by year, allowing us to work with annual values instead of monthly layers.

```
crp = crop(r,box)
filtered_values <- crp[[(1369:1440)]]
annual_raster <- tapp(filtered_values, index = as.numeric(format(time(filtered_values), "%Y")), fun = me
```

Task 3

Setting seed and create 6 random numbers.

```
set.seed(702)
random_numbers <- runif(6, min = 0, max = 1)
random_numbers
```

```
## [1] 0.55636888 0.43993669 0.12869645 0.26272641 0.08981997 0.76960369
```

We took the aggregated raster and multiplied each of the six annual layers by a unique set of randomly generated numbers, ensuring reproducibility by using a specific seed.

```
values(annual_raster[[1]]) <- na.omit(values(annual_raster[[1]]))
values(annual_raster[[2]]) <- na.omit(values(annual_raster[[2]]))
values(annual_raster[[3]]) <- na.omit(values(annual_raster[[3]]))
values(annual_raster[[4]]) <- na.omit(values(annual_raster[[4]]))
values(annual_raster[[5]]) <- na.omit(values(annual_raster[[5]]))
values(annual_raster[[6]]) <- na.omit(values(annual_raster[[6]]))

annual_raster[[1]] <- annual_raster[[1]] * random_numbers[1]
annual_raster[[2]] <- annual_raster[[2]] * random_numbers[2]
annual_raster[[3]] <- annual_raster[[3]] * random_numbers[3]
annual_raster[[4]] <- annual_raster[[4]] * random_numbers[4]
annual_raster[[5]] <- annual_raster[[5]] * random_numbers[5]
annual_raster[[6]]<- annual_raster[[6]] * random_numbers[6]
```

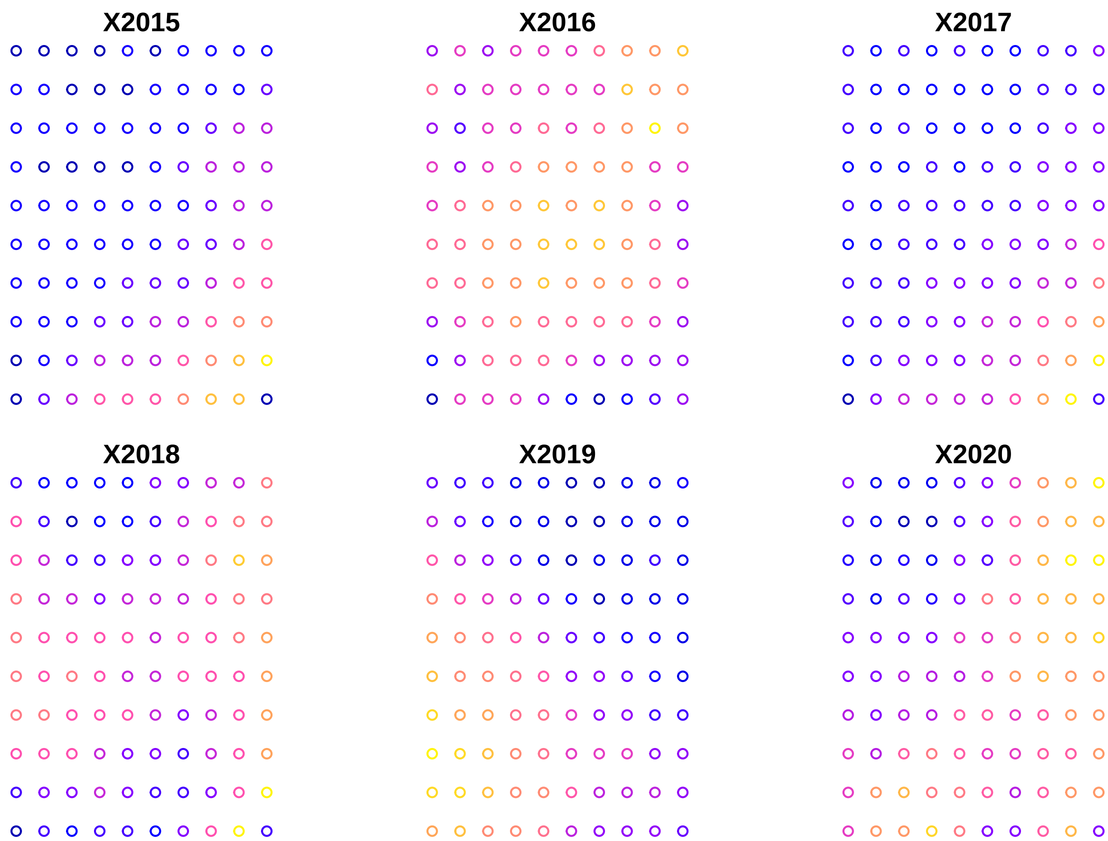
Task 4

Then, we extract points from raster as transform them to sf object.

```
points <- as.points(annual_raster)
points
```

```
## class       : SpatVector
## geometry    : points
## dimensions  : 100, 6  (geometries, attributes)
## extent      : 18.75, 23.25, 41.75, 46.25  (xmin, xmax, ymin, ymax)
## coord. ref. : lon/lat WGS 84 (CRS84) (OGC:CRS84)
## names       :   X2015   X2016    X2017    X2018    X2019   X2020
## type        :   <num>   <num>    <num>    <num>    <num>   <num>
## values      : -0.2124 0.07772 -0.05227 -0.08664 -0.03772 -0.2635
##                 -0.227 0.08014 -0.06081 -0.09788 -0.04348 -0.3365
##               -0.2049 0.07538 -0.05769 -0.09762 -0.04475 -0.3239
```

```
points_sf <- st_as_sf(points)
plot(points_sf)
```



Task 5

Since we have SPEI data in raster format, to facilitate analysis, we need to extract and convert it into polygons. This allows us to view the data in table format and visualize it on plots. We start by aligning the SPEI data with the shp (country shapefile) and shp3 (NUTS 3 regions) shapefiles. To retain the spatial geometry, we merge the extracted values back with shp and shp3, ensuring a seamless integration of data and geometry

shp

```
yearly_data <- list()
years <- 2015:2020
```

```r
for (i in 1:length(years)) {
  year <- years[i]
  raster_layer <- annual_raster[[i]]

  transformed_shp <- st_transform(shp, crs = st_crs(raster_layer))

  extracted_data <- terra::extract(raster_layer, vect(transformed_shp), fun = mean, na.rm = TRUE)

  col_name <- paste0("SPEI_", year)
  extracted_data <- extracted_data %>% rename(!!col_name := paste0("X", year))

  yearly_data[[as.character(year)]] <- cbind(transformed_shp, extracted_data)
}

final_shp <- yearly_data[["2015"]]

for (i in 2:length(years)) {
  year <- years[i]
  col_name <- paste0("SPEI_", year)
  final_shp <- final_shp %>% mutate(!!col_name := yearly_data[[as.character(year)]][[col_name]])
}
```

shp3

```r
years <- 2015:2020
shp_list <- list()

for (i in seq_along(years)) {
  year <- years[i]

  shp_transformed <- st_transform(shp, crs = st_crs(annual_raster[[i]]))

  spei_data <- terra::extract(annual_raster[[i]], vect(shp), fun = mean, na.rm = TRUE)

  spei_data <- spei_data %>% rename(!!paste0("SPEI_", year) := !!paste0("X", year))
  shp_list[[i]] <- cbind(shp_transformed, spei_data)
}

final_shp <- shp_list[[1]]
for (i in 2:length(years)) {
  final_shp <- final_shp %>% mutate(!!paste0("SPEI_", years[i]) := shp_list[[i]][[paste0("SPEI_", years
}

shp3_list <- list()
for (i in seq_along(years)) {
  year <- years[i]

  shp3_transformed <- st_transform(shp3, crs = st_crs(annual_raster[[i]]))

  spei_data <- terra::extract(annual_raster[[i]], vect(shp3), fun = mean, na.rm = TRUE)
```

```r
  spei_data <- spei_data %>% rename(!!paste0("SPEI_", year) := !!paste0("X", year))
  shp3_list[[i]] <- cbind(shp3_transformed, spei_data)
}

final <- shp3_list[[1]]
for (i in 2:length(years)) {
  final <- final %>% mutate(!!paste0("SPEI_", years[i]) := shp3_list[[i]][[paste0("SPEI_", years[i])]])
}
```

Calculating the mean values for the years 2015 and 2020

```r
mean(final$SPEI_2015)
```

```
## [1] -0.1568477
```

```r
mean(final$SPEI_2020)
```

```
## [1] -0.2515367
```

These values indicate a decrease in the average SPEI (Standardized Precipitation-Evapotranspiration Index) from 2015 to 2020, suggesting a trend toward drier conditions or increased water stress over this period. This shift in SPEI may reflect broader climatic changes impacting the region.

Task 6

Now we go to another shape file that contains data of Aqueduct

```r
aq_shp <- st_read("C:/Users/jelce/Documents/Study/GIS/Exam/Aqueduct_baseline.shp")
```

```
## Reading layer `Aqueduct_baseline' from data source
##   `C:\Users\jelce\Documents\Study\GIS\Exam\Aqueduct_baseline.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 382 features and 237 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: 13.49011 ymin: 42.22762 xmax: 22.9788 ymax: 48.58638
## Geodetic CRS:  WGS 84
```

Task 7

```r
set.seed(702)
length(aq_shp$bws_raw)
```

```
## [1] 382
```

```r
water_stress_2015 <- aq_shp$bws_raw
aq_shp$water_stress_2015 <- water_stress_2015
multipliers <- matrix(rnorm(5 * 382, mean = 1, sd = 1), nrow = 382, ncol = 5)
water_stress_ts <- water_stress_2015 * multipliers
```

```r
aq_shp$water_stress_2016 <- water_stress_ts[, 1]
aq_shp$water_stress_2017 <- water_stress_ts[, 2]
aq_shp$water_stress_2018 <- water_stress_ts[, 3]
aq_shp$water_stress_2019 <- water_stress_ts[, 4]
aq_shp$water_stress_2020 <- water_stress_ts[, 5]
```

Task 8

```r
columns_to_select <- c("geometry", paste0("water_stress_", 2015:2020))
aq_shp_selected <- aq_shp %>%
  dplyr::select(all_of(columns_to_select))
```

Since we have acquired the data as a shapefile, we can convert it into raster format, which is often more suitable for spatial analysis and visualization. By converting to a raster, we can efficiently analyze spatial variations and perform calculations on a pixel-by-pixel basis.

To handle multiple years of data, a loop was employed to automate the rasterization process for each year, ensuring consistency and saving time compared to manual processing.

```r
r.template <- rast(ext = ext(shp3), resolution = 0.01, crs = crs(shp3))
water_stress_rasters <- list()
for (year in 2015:2020) {
  attribute_name <- paste0("water_stress_", year)
  water_stress_rasters[[as.character(year)]] <- rasterize(vect(aq_shp), r.template, field = attribute_na
}
```

Then we extract data from raster

```r
serbia_values <- list()
nuts3_values <- list()

for (year in 2015:2020) {
  serbia_values[[as.character(year)]] <- terra::extract(water_stress_rasters[[as.character(year)]], vect

  nuts3_values[[as.character(year)]] <- terra::extract(water_stress_rasters[[as.character(year)]], vect
  }
```

The average water stress level across Serbian NUTS 3 districts in 2020 shows a slight increase compared to the overall simple average. This indicates that, on a district level, some areas may be experiencing heightened water stress, which could reflect localized environmental or resource management challenges. This nuanced difference between the district-level and overall averages highlights the importance of regional analysis in understanding water stress patterns, as aggregate figures may overlook these localized variations.

```
serbia_avg_2020 <- serbia_values[["2020"]]
serbia_avg_2020$water_stress_2020
```

```
## [1] 0.4107806
```

```
nuts3_avg_2020 <- mean(nuts3_values[["2020"]][, 2], na.rm = TRUE)
nuts3_avg_2020
```

```
## [1] 0.4550146
```

Task 9

Load the raster of population density for 2015.

```
pop_den = rast("C:/Users/jelce/Documents/Study/GIS/Exam/gpw_v4_population_density_rev11_2015_15_min.asc
```

Cropping it and calculating the average.

```
box <- st_bbox(shp)
box <- box+c(-.5,-.5,.5,.5)
den_crop = crop(pop_den,box)
serbia_density_values <- terra::extract(den_crop, vect(shp3), fun = mean, na.rm = TRUE)
mean(serbia_density_values$gpw_v4_population_density_rev11_2015_15_min)
```

```
## [1] 97.55682
```

In this analysis, we observe that the initial mean population density for Serbia's NUTS 3 regions, based on the original cropped data, is approximately 97.56. After resampling the data using the "bilinear" method, the mean value increased slightly to 106.81. This increase indicates that resampling, especially with a method like bilinear interpolation, can smooth and alter the data by averaging neighboring cells, which may lead to slightly higher values. Such differences highlight the impact of resampling techniques on data interpretation, especially when analyzing population density or other spatial data.

```
resampled_density <- resample(den_crop, filtered_values, method = "bilinear")
serbia_resampled_density_values <- terra::extract(resampled_density, vect(shp3), fun = mean, na.rm = TRU
mean(serbia_resampled_density_values$gpw_v4_population_density_rev11_2015_15_min)
```

```
## [1] 106.8086
```

Task 10

When analyzing GDP data, using population-weighted climate variables can be insightful, especially in densely populated regions where climate conditions may directly impact productivity and economic output.

For Agricultural GVA or yields, population weighting is generally less critical, as agricultural productivity is more influenced by land quality and environmental conditions. However, population-weighted data might capture indirect effects in densely inhabited areas, where human activity influences agricultural practices.

In the case of Industrial GVA, population-weighted climate data may be relevant if climate significantly affects labor productivity or operational conditions in heavily populated industrial zones.

Task 11

Reducing the SPEI resolution to match the finer population density resolution (0.25, 0.25) is preferable for creating a population-weighted SPEI index. This approach retains the spatial detail of the population data, enabling more accurate insights into how climate impacts may vary across densely populated areas. If we instead aggregated population data to the coarser SPEI resolution (1,1), we would lose important localized information, limiting the precision of our analysis. # Task 12 Loading GVA data.

```
library(readxl)
library(dplyr)
gva <- read_xlsx("serbian_gva_sector_a.xlsx", sheet = 2) %>% as.data.frame()
gva <- gva %>% dplyr::select("TERRITORY_ID", "LEVEL_ID", "NAME_HTML","VERSIONS", "UNIT", "SECTOR", conta
```

We rename the columns appropriately to clearly indicate that they refer to GVA values. Then, we select only the necessary columns and remove any duplicate columns from shp3 before merging. Finally, we merge the files, ensuring a clean and well-structured dataset for analysis.

```
gva <- gva %>%
  rename(
    gva_2015 = `2015`,
    gva_2016 = `2016`,
    gva_2017 = `2017`,
    gva_2018 = `2018`,
    gva_2019 = `2019`,
    gva_2020 = `2020`,
    NUTS_ID = TERRITORY_ID
  )
gva = gva %>% select(NUTS_ID, gva_2015, gva_2016, gva_2017, gva_2018, gva_2019, gva_2020)
shp3 = final %>% select(-FID)
final_sf = shp3 %>% left_join(gva, by = "NUTS_ID")
```

Task 13

To create the final sf object, we focus primarily on the water stress data, as the GVA data has already been incorporated. We retrieve the values for each year, previously stored as separate objects, and consolidate the data for all years into a single object. Finally, we merge this consolidated water stress data with the rest of our dataset to complete the final sf object.

```
water_stress_2015 <- nuts3_values[[1]]
water_stress_2016 <- nuts3_values[[2]]
water_stress_2017 <- nuts3_values[[3]]
water_stress_2018 <- nuts3_values[[4]]
water_stress_2019 <- nuts3_values[[5]]
water_stress_2020 <- nuts3_values[[6]]

# Ensure only the ID column and the year-specific water stress columns are selected
water <- water_stress_2015 %>%
  left_join(water_stress_2016 %>% select(ID, water_stress_2016), by = "ID") %>%
  left_join(water_stress_2017 %>% select(ID, water_stress_2017), by = "ID") %>%
  left_join(water_stress_2018 %>% select(ID, water_stress_2018), by = "ID") %>%
  left_join(water_stress_2019 %>% select(ID, water_stress_2019), by = "ID") %>%
```

```
  left_join(water_stress_2020 %>% select(ID, water_stress_2020), by = "ID")
final_sf = final_sf %>% left_join(water, by = "ID")
```

Task 14

```
gva_serbia = gva %>% summarise (gva_2015 = mean(gva_2015), gva_2016 = mean(gva_2016), gva_2017 = mean(g
```

To plot the time series of SPEI and water stress for Serbia from 2015 to 2020, we need to calculate the average for each year, as the object contains only a single observation representing Serbia.

```
water_avg <- colMeans(water[, -1], na.rm = TRUE)
water_avg_df <- as.data.frame(t(water_avg))
final_serbia = cbind(final_shp, water_avg_df)
```

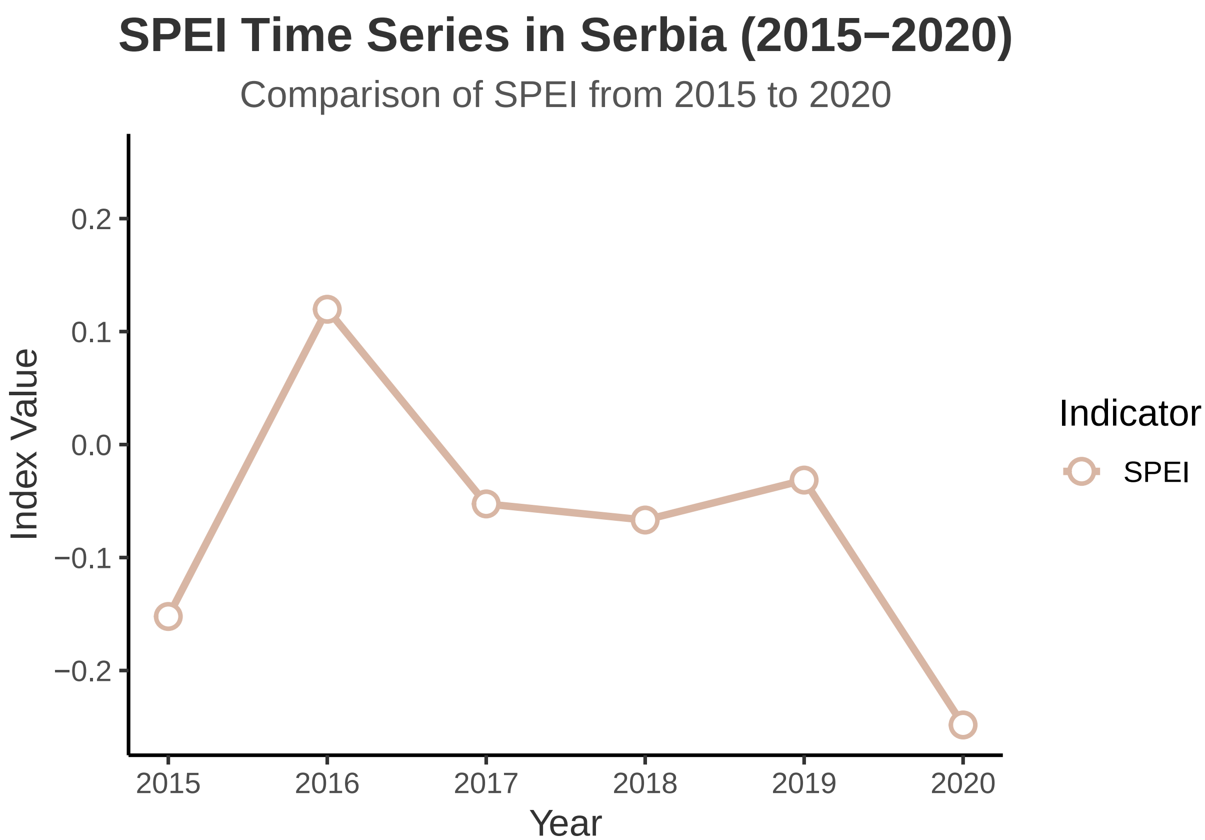
At this stage, our dataset contains separate columns for each specific year, but to visualize the trends over time and compare multiple variables on the same graph, we need to reshape the data into a 'long' format. This transformation allows us to plot all years dynamically in a single time series. To achieve this, we use the pivot_longer function, which consolidates the yearly columns into a unified 'year' column. However, during this process, we encountered an issue where the year values appeared as NA instead of the correct years, while the water stress values themselves were intact. To resolve this, we manually reassigned the correct year values.

```
library(ggplot2)
library(dplyr)
library(tidyr)

data_no_geom <- final_serbia %>%
  st_set_geometry(NULL) %>%
  select(matches("SPEI|water_stress_"))
data_long <- data_no_geom %>%
  pivot_longer(cols = everything(),
               names_to = c("variable", "year"),
               names_sep = "_") %>%
  mutate(year = as.numeric(year))
water_only <- data_long %>%
  filter(variable == "water")
w <- water_only %>%
  mutate(year = rep(2015:2020, each = 1))
```

First, we do it for SPEI data.

```
library(ggplot2)

ggplot(data_long, aes(x = year, y = value, color = variable, group = variable)) +
  geom_line(size = 1.3) +
  geom_point(size = 3.5, shape = 21, fill = "white", stroke = 1.2) +
  ylim(-0.25, 0.25) +
  scale_color_manual(values = c("SPEI" = "#d8b6a4", "water_stress" = "#c9a6a6")) +
  labs(
    title = "SPEI Time Series in Serbia (2015-2020)",
```

```r
    subtitle = "Comparison of SPEI from 2015 to 2020",
    x = "Year",
    y = "Index Value",
    color = "Indicator"
  ) +
  theme_classic(base_size = 14, base_family = "Helvetica") +
  theme(
    plot.title = element_text(size = 18, face = "bold", hjust = 0.5, color = "#333333"),
    plot.subtitle = element_text(size = 14, hjust = 0.5, color = "#555555"),
    axis.title.x = element_text(size = 14, color = "#333333"),
    axis.title.y = element_text(size = 14, color = "#333333")
  )
```



This graph displays the Standardized Precipitation Evapotranspiration Index (SPEI) time series in Serbia from 2015 to 2020, showing fluctuations in the index value over the specified period. Each data point represents the average SPEI value for a given year. The graph indicates a peak in SPEI around 2016, followed by a steady decline, with a notable drop in 2020, suggesting an increase in arid conditions in recent years. The light brown color scheme provides a subdued aesthetic, with the larger marker points making each year's value easy to identify. The consistent, clean layout ensures clarity, allowing for straightforward interpretation of SPEI trends over time.

Then we plot water stress level in Serbia.

```r
ggplot(data = w, aes(x = year, y = value, group = 1)) +
  geom_line(color = "#4a90e2", size = 1.3) +  # Cool blue color
  geom_point(color = "#4a90e2", size = 3, shape = 21, fill = "white", stroke = 1) +
  labs(
    title = "Water Stress Time Series in Serbia (2015-2020)",
```

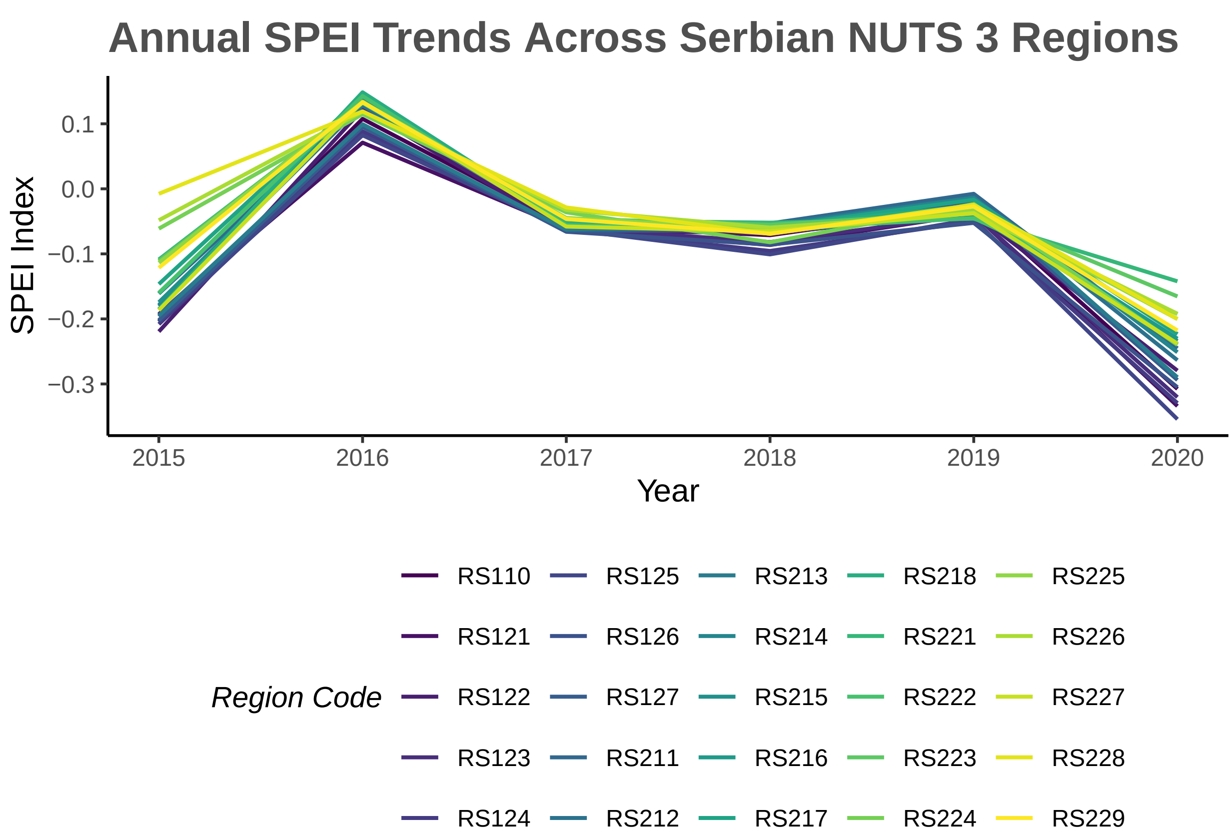
```
  subtitle = "Annual Water Stress Index from 2015 to 2020",
  x = "Year",
  y = "Water Stress Value"
) +
theme_classic(base_size = 14, base_family = "Helvetica") +
theme(
  plot.title = element_text(size = 18, face = "bold", hjust = 0.5, color = "#333333"),
  plot.subtitle = element_text(size = 14, hjust = 0.5, color = "#555555"),
  axis.title.x = element_text(size = 14, color = "#333333"),
  axis.title.y = element_text(size = 14, color = "#333333")
)
```

This graph represents the Water Stress Index time series in Serbia from 2015 to 2020, showing the annual fluctuations in water stress. The most significant peak occurs in 2016, where water stress reached its highest level in this time frame. Following 2016, there is a sharp decline in water stress values, reaching a low between 2017 and 2018. However, from 2018 onwards, there is a gradual upward trend, indicating increasing water stress levels. The vivid blue line and markers highlight the changing trend across the years, and the clear layout emphasizes the progression of water stress, making it easy to observe the variability over time. This complements the SPEI graph by providing a contrasting yet related view on environmental stress factors in Serbia.

Task 15

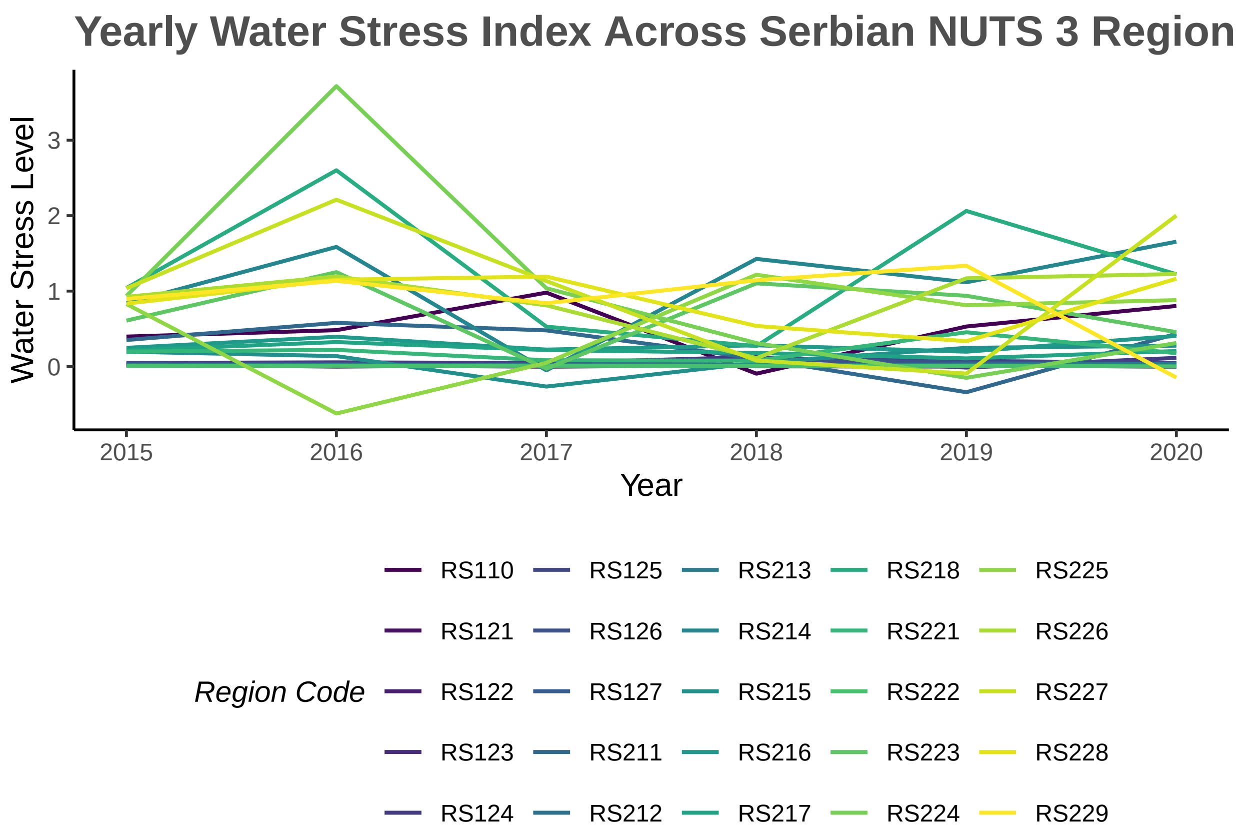
Here's the plot that visualizes the time series for each of the 25 Serbian NUTS 3 regions. Each line represents a region, allowing us to see individual trends in the data over time.

```r
spei_data_long <- final_sf %>%
  select(region_code = NUTS_ID, starts_with("SPEI")) %>%
  pivot_longer(cols = starts_with("SPEI"),
               names_to = "year_col",
               values_to = "spei_index") %>%
  mutate(year = as.numeric(sub("SPEI_", "", year_col)))

ggplot(spei_data_long, aes(x = year, y = spei_index, color = region_code, group = region_code)) +
  geom_line(size = 0.7) +
  labs(
    title = "Annual SPEI Trends Across Serbian NUTS 3 Regions",
    x = "Year",
    y = "SPEI Index"
  ) +
  theme_classic() +
  theme(
    legend.position = "bottom",
    legend.title = element_text(face = "italic"),
    plot.title = element_text(size = 16, face = "bold", color = "#4f4f4f"),
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12)
  ) +
  scale_color_viridis_d(name = "Region Code")
```



This plot shows the SPEI (Standardized Precipitation-Evapotranspiration Index) time series for each of the 25 Serbian NUTS 3 regions from 2015 to 2020. Each line represents the SPEI values for a specific region over time, illustrating trends in drought or wetness conditions. The graph reveals that most regions follow a similar pattern, with peaks around 2016 and declines around 2020. However, subtle variations among regions indicate some regional differences in response to climatic factors. The use of different colors for each region helps distinguish between them, though it may be challenging to track specific regions due to overlapping lines. This visualization provides an overview of SPEI trends across regions, helping to identify periods of stress or relief across Serbian NUTS 3 areas.

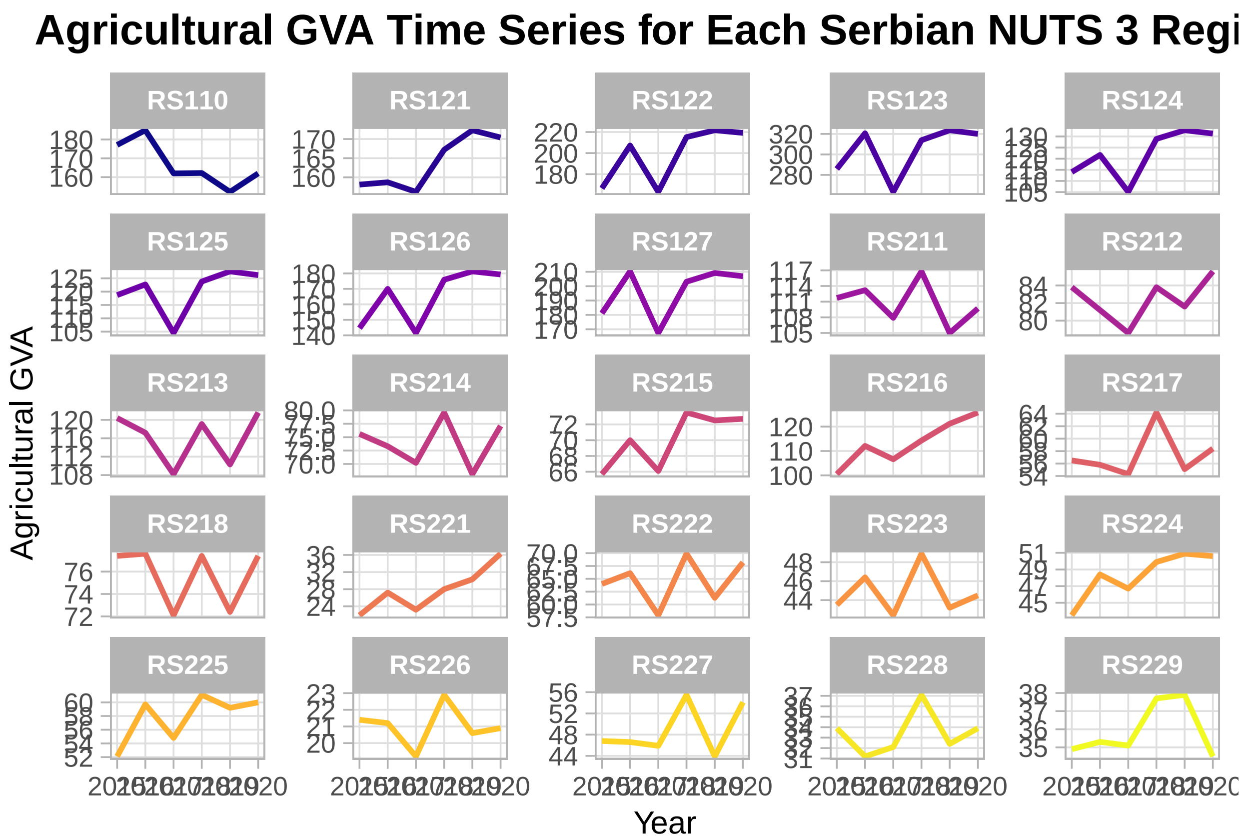
```
water_stress_long <- final_sf %>%
  select(region_id = NUTS_ID, starts_with("water_stress")) %>%
  pivot_longer(cols = starts_with("water_stress"),
               names_to = "time",
               values_to = "stress_value") %>%
  mutate(year = as.numeric(sub("water_stress_", "", time)))
ggplot(water_stress_long, aes(x = year, y = stress_value, color = region_id, group = region_id)) +
  geom_line(size = 0.7) +
  labs(
    title = "Yearly Water Stress Index Across Serbian NUTS 3 Regions",
    x = "Year",
    y = "Water Stress Level"
  ) +
  theme_classic() +
  theme(
    legend.position = "bottom",
    legend.title = element_text(face = "italic"),
    plot.title = element_text(size = 16, face = "bold", color = "#4f4f4f"),
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12)
  ) +
  scale_color_viridis_d(name = "Region Code")
```



This graph presents the water stress index for each Serbian NUTS 3 region from 2015 to 2020. Each line represents a region, identified by its unique code (e.g., RS110, RS121), showing the evolution of water stress levels over time. The plot highlights regional differences in water stress, with some areas experiencing noticeable variability and others remaining relatively stable. The diverse color scheme helps differentiate regions, though further adjustments to axis labels could enhance readability. This visualization provides a comprehensive overview of regional water stress patterns and facilitates easy comparison across Serbia.

```r
gva_long <- final_sf %>%
  select(NUTS_ID, starts_with("gva")) %>%
  pivot_longer(cols = starts_with("gva"), names_to = "year", values_to = "gva_value") %>%
  mutate(year = as.numeric(gsub("gva_", "", year)))

ggplot(gva_long, aes(x = year, y = gva_value, color = NUTS_ID)) +
  geom_line(size = 1) +
  facet_wrap(~ NUTS_ID, scales = "free_y", ncol = 5) +
  labs(title = "Agricultural GVA Time Series for Each Serbian NUTS 3 Region",
       x = "Year", y = "Agricultural GVA") +
  theme_light(base_size = 14, base_family = "Helvetica") +
  theme(
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
    axis.title.x = element_text(size = 12),
    axis.title.y = element_text(size = 12),
    axis.text = element_text(size = 10),
    strip.text = element_text(size = 10, face = "bold"),
    legend.position = "none",
    panel.grid.minor = element_blank()
  ) +
```

```r
  scale_color_viridis_d(option = "C")
```



This graph illustrates the agricultural GVA time series for each Serbian NUTS 3 region from 2015 to 2020. Each panel represents a region, identified by its NUTS code (e.g., RS110, RS121), and shows the annual changes in agricultural gross value added (GVA).

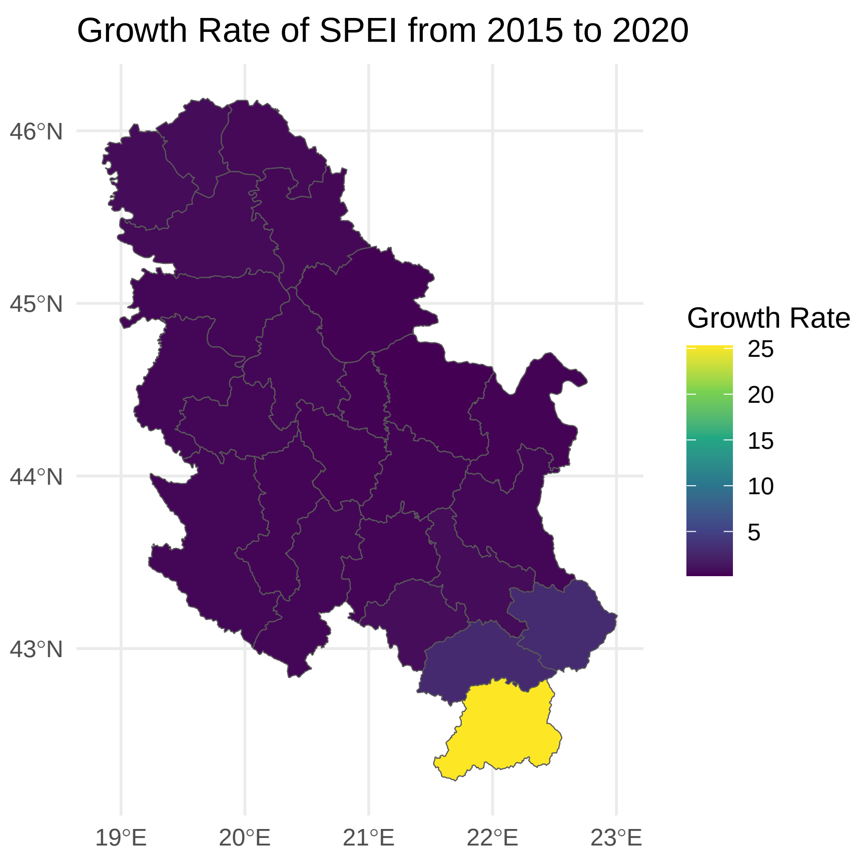
The variations across regions are evident, with some regions showing significant fluctuations, while others demonstrate more stable trends over the years. The color coding of lines by region helps in distinguishing between them, while the small multiples layout allows for easy cross-regional comparison. However, the x-axis labels appear crowded; adjusting them could improve readability. Additionally, labeling the y-axis as "Agricultural GVA" instead of "Water Stress Index" would better reflect the data being presented.

Task 16

We calculated the growth rate of SPEI (Standardized Precipitation Evapotranspiration Index) from 2015 to 2020 using the formula provided. However, we encountered an outlier issue with the region "Пчињска област," where the growth rate was so high that it dominated the scale of the graph, making it difficult to visualize differences between this district and the others. This extreme value affects the overall interpretation and visualization, as it compresses the range for other regions, limiting our ability to discern regional variations in SPEI growth effectively.

```r
growth = final_sf %>%
  mutate(speigr = (SPEI_2020 - SPEI_2015) / SPEI_2015) %>%
  select(NUTS_NAME, speigr, geometry)

ggplot(data = growth) +
```

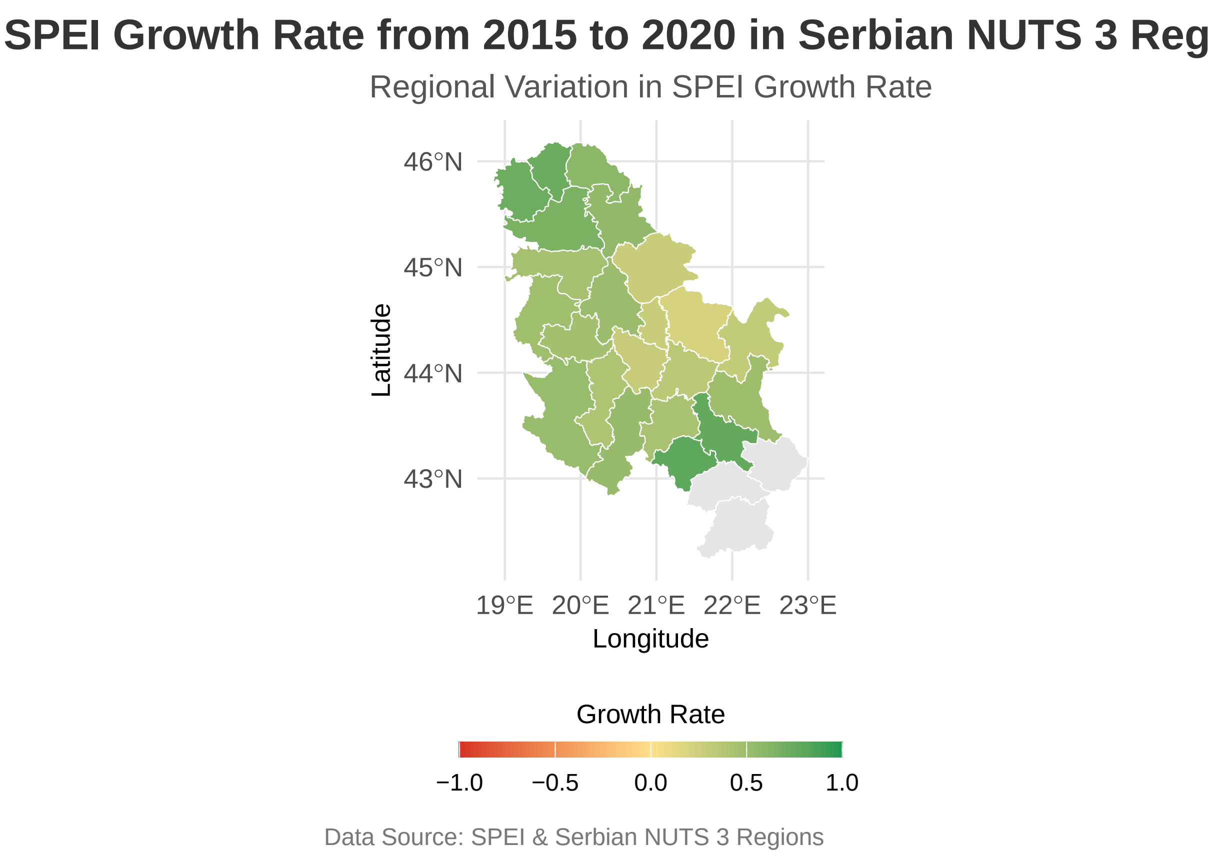
```r
  geom_sf(aes(fill = speigr)) +
  scale_fill_viridis_c() +
  labs(
    title = "Growth Rate of SPEI from 2015 to 2020",
    fill = "Growth Rate"
  ) +
  theme_minimal()
```



To address the outlier issue and better visualize regional diversity in terms of SPEI growth rate, we applied filters to focus on regions within a manageable growth rate range.

```r
library(ggplot2)
library(sf)

ggplot(data = growth) +
  geom_sf(aes(fill = speigr), color = "white", size = 0.2) +
  scale_fill_gradientn(
    colors = c("#d73027", "#fee08b", "#1a9850"),  # Custom gradient
    na.value = "grey90",
    limits = c(-1, 1),
    name = "Growth Rate"
  ) +
  labs(
    title = "SPEI Growth Rate from 2015 to 2020 in Serbian NUTS 3 Regions",
    subtitle = "Regional Variation in SPEI Growth Rate",
    caption = "Data Source: SPEI & Serbian NUTS 3 Regions",
```

```
    x = "Longitude",
    y = "Latitude"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5, color = "#333333"),
    plot.subtitle = element_text(size = 12, hjust = 0.5, color = "#555555"),
    plot.caption = element_text(size = 9, color = "#777777", hjust = 1),
    legend.position = "bottom",
    legend.title = element_text(size = 10),
    legend.text = element_text(size = 9),
    panel.grid.major = element_line(color = "#e5e5e5", size = 0.4),
    panel.grid.minor = element_blank(),
    axis.title.x = element_text(size = 10),
    axis.title.y = element_text(size = 10)
  ) +
  guides(
    fill = guide_colorbar(
      barwidth = 10, barheight = 0.4,
      title.position = "top", title.hjust = 0.5
    )
  )
```



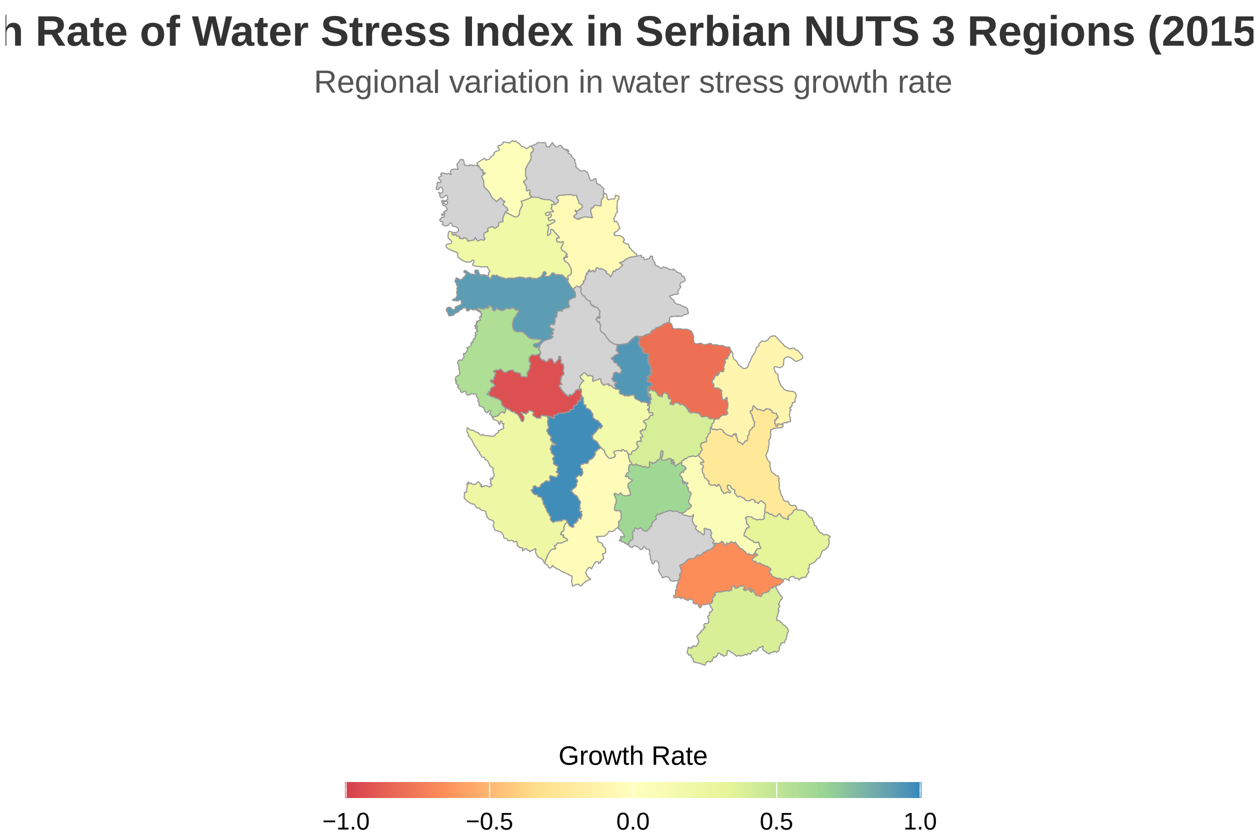
The map above illustrates the SPEI (Standardized Precipitation-Evapotranspiration Index) growth rate across Serbian NUTS 3 regions from 2015 to 2020. Green shades represent positive growth, where regions experienced an increase in SPEI values, indicating an improvement in water balance conditions over this period. In contrast, the light yellow-green tones suggest minimal or neutral growth.

Notably, the central part of Serbia shows the lowest growth rates, with relatively stable or only slightly improved SPEI values. This suggests that, compared to other regions, central Serbia experienced less change in water balance conditions from 2015 to 2020. The visual distribution of colors across the map provides a quick reference to understand regional variations in water stress response over time, with the highest positive growth observed in the northern and southern areas.

Task 17

Then, we do the same for water stress data, but chnaging CRS to 32634.

```
final_sf <- final_sf %>%
  mutate(wsgr = (water_stress_2020 - water_stress_2015) / water_stress_2015)
final_sf <- st_transform(final_sf, crs = 32634)

ggplot(data = final_sf) +
  geom_sf(aes(fill = wsgr), color = "grey60", size = 0.3) +
  scale_fill_distiller(
    palette = "Spectral", direction = 1, na.value = "lightgrey",
    limits = c(-1, 1)
  ) +
  labs(
    title = "Growth Rate of Water Stress Index in Serbian NUTS 3 Regions (2015-2020)",
    subtitle = "Regional variation in water stress growth rate",
    fill = "Growth Rate"
  ) +
  theme_minimal(base_size = 11) +
  theme(
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5, color = "#333333"),
    plot.subtitle = element_text(size = 12, hjust = 0.5, color = "#555555"),
       legend.position = "bottom",
    legend.title = element_text(size = 10),
    legend.text = element_text(size = 9),
    panel.grid.major = element_blank(),
    axis.title = element_blank(),
    axis.text = element_blank(),
    axis.ticks = element_blank()
  ) +
  guides(
    fill = guide_colorbar(
      barwidth = 15, barheight = 0.4,
      title.position = "top", title.hjust = 0.5
    )
  )
```



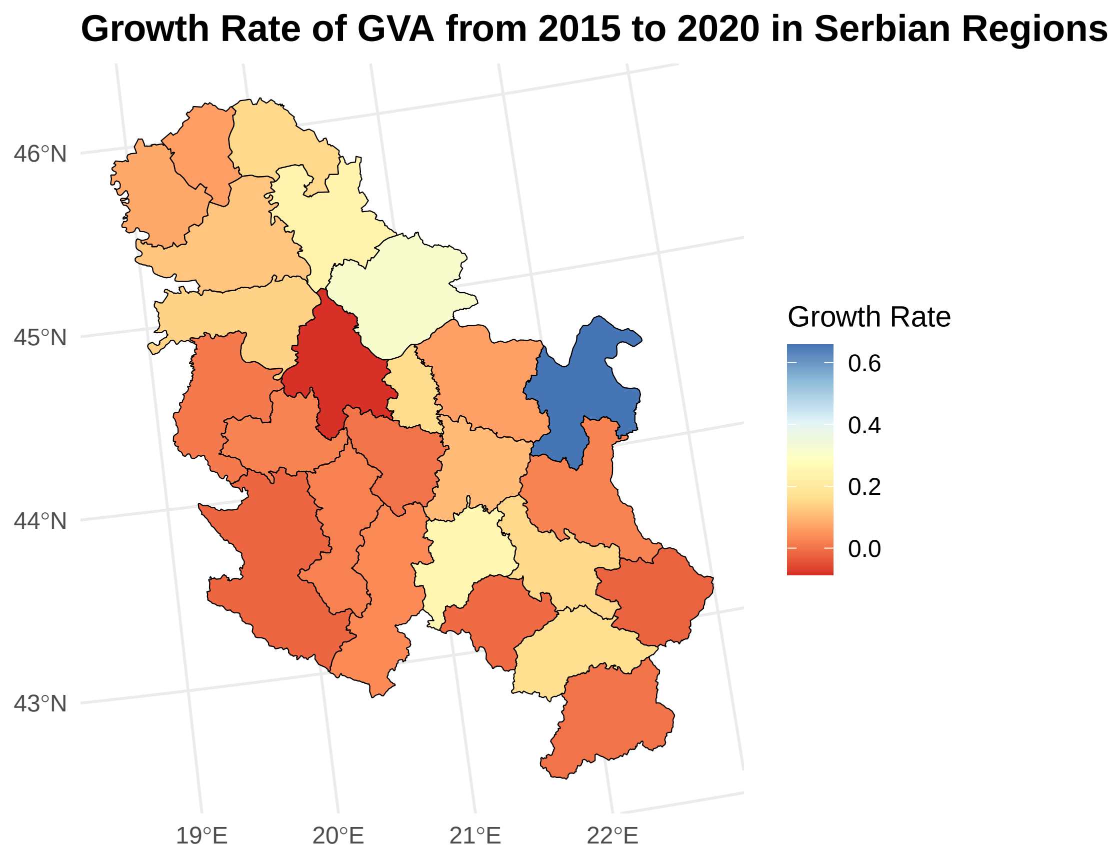
The map above illustrates the growth rate of the Water Stress Index across Serbian NUTS 3 regions from 2015 to 2020, using CRS 32634 to provide an accurate spatial representation. The color gradient reflects regional changes in water stress over this period: blue tones represent positive growth, indicating a decrease in water stress levels, while red tones signify negative growth, suggesting an increase in water stress levels.

This map reveals notable regional disparities. Central and southern Serbia exhibit higher levels of negative growth (in red and orange), indicating worsening water stress conditions in these areas. In contrast, regions in the north and southeast, represented by blue tones, show an improvement, with water stress levels decreasing over time. The most pronounced contrasts appear in central Serbia, where the increase in water stress is the highest, highlighting potential areas for targeted water management strategies.

Now we plot the growth rate Of GVA from 2015 to 2020.

```
final_sf <- final_sf %>%
  mutate(agrgvagr = (gva_2020 - gva_2015) / gva_2015)
final_sf <- st_transform(final_sf, crs = 3035)

ggplot(data = final_sf) +
  geom_sf(aes(fill = agrgvagr), color = "black", size = 0.2) +
  scale_fill_distiller(palette = "RdYlBu", direction = 1, na.value = "grey90") +
  labs(
    title = "Growth Rate of GVA from 2015 to 2020 in Serbian Regions",
    fill = "Growth Rate"
  ) +
  theme_minimal() +
  theme(
    plot.title = element_text(size = 14, face = "bold"),
```

```
  legend.position = "right"
)
```

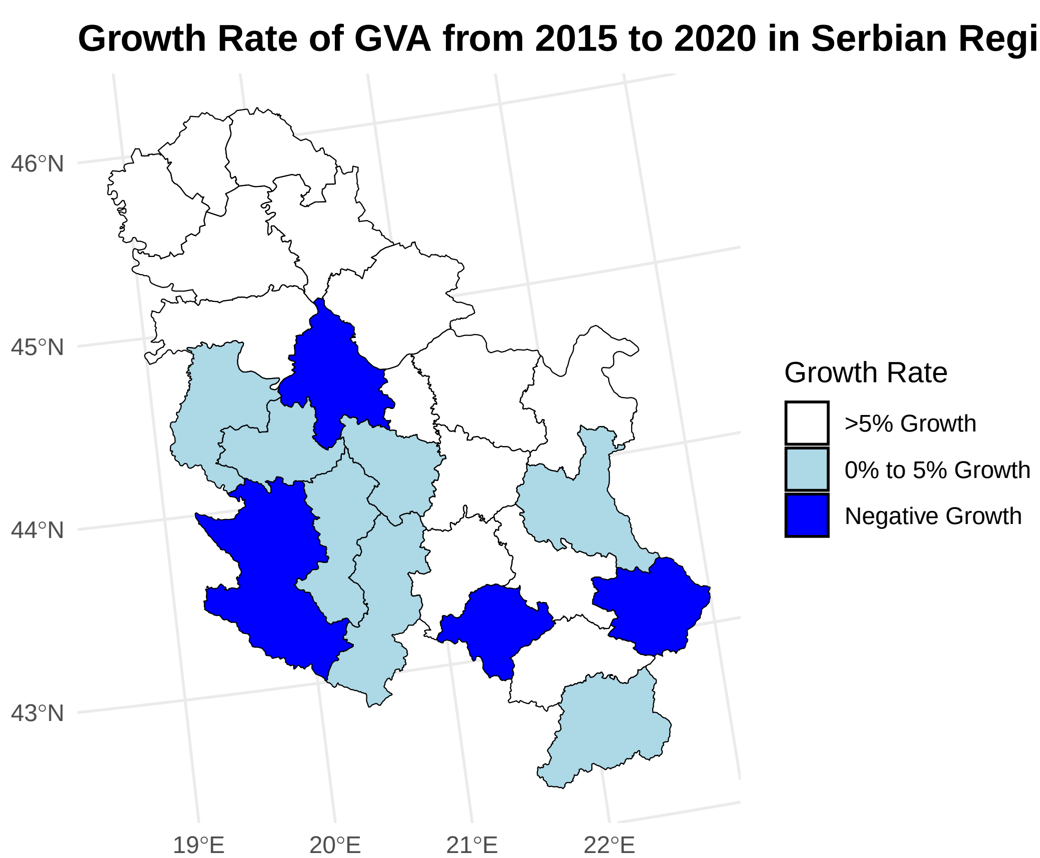


The map illustrates the growth rate of Agricultural Gross Value Added (GVA) across Serbian NUTS 3 regions from 2015 to 2020, displayed in CRS 3035 for precise spatial representation. The color gradient captures different growth levels, with blue shades indicating positive growth and red shades signaling stagnation or decline in agricultural GVA.

Notably, several regions in the north and northeast of Serbia, depicted in blue, demonstrated substantial growth in agricultural productivity, possibly reflecting advancements in agricultural practices, increased investments, or favorable climate conditions. In contrast, central and southern regions are predominantly colored in red and orange, indicating minimal to no growth, and in some cases, a decrease in agricultural GVA during this period. This distribution suggests regional disparities in agricultural performance, potentially influenced by factors such as soil quality, climate variability, and access to essential resources and infrastructure.

```
final_sf <- final_sf %>%
  mutate(agrgvagr = (gva_2020 - gva_2015) / gva_2015)
final_sf <- st_transform(final_sf, crs = 3035)

final_sf <- final_sf %>%
  mutate(
    growth_category = case_when(
      agrgvagr < 0 ~ "Negative Growth",
      agrgvagr >= 0 & agrgvagr <= 0.05 ~ "0% to 5% Growth",
      agrgvagr > 0.05 ~ ">5% Growth"
    )
  )
```

```r
ggplot(data = final_sf) +
  geom_sf(aes(fill = growth_category), color = "black", size = 0.2) +
  scale_fill_manual(
    values = c(
      "Negative Growth" = "blue",
      "0% to 5% Growth" = "lightblue",
      ">5% Growth" = "white"
    ),
    na.value = "grey90"
  ) +
  labs(
    title = "Growth Rate of GVA from 2015 to 2020 in Serbian Regions",
    fill = "Growth Rate"
  ) +
  theme_minimal() +
  theme(
    plot.title = element_text(size = 14, face = "bold"),
    legend.position = "right"
  )
```



We applied this filtering to make regional disparities in Agricultural GVA growth in Serbia from 2015 to 2020 more visually clear and easier to interpret. The map reveals that northern and northeastern regions experienced higher growth, likely benefiting from favorable conditions, investments, or advancements in agriculture, reflecting increased productivity. In contrast, central and southern areas show lower or negative growth, potentially due to resource limitations, unfavorable climate, or structural issues. This variation underscores the need for targeted policies to support balanced agricultural development and address regional inequalities across Serbia.

Task 19

To assess whether the GDP per capita in a region could be influenced by the lagged investment in a neighboring region, we can utilize spatial econometric methods to account for spatial dependencies among regions. The following approach outlines a structured methodology that aligns with our analysis of regional data in this project:

1. Define Spatial Relationships
   Construct a spatial weights matrix that defines relationships between neighboring NUTS 3 regions. This matrix will typically be based on geographic proximity (e.g., neighboring borders or distances) and is essential for capturing the influence of investment from one region on its neighbors.
2. Apply Spatial Econometric Models
   Use spatial econometric models such as the Spatial Lag Model (SLM) or Spatial Durbin Model (SDM). These models allow us to model GDP per capita in a region as a function of its own lagged investment and the lagged investment in neighboring regions, as defined by the spatial weights matrix.
3. Incorporate Temporal Lags
   Since our dataset has a monthly frequency, we can introduce temporal lags to capture the impact of investments from previous months on the current GDP per capita. This addition provides insights into both immediate and delayed effects.
4. Assess Spatial Autocorrelation
   Check for spatial autocorrelation in GDP per capita using Moran's I statistic. A significant positive Moran's I value would indicate that GDP per capita in one region is indeed spatially correlated with neighboring regions, supporting the notion of regional interdependence.

This approach enables us to explore the spatial and temporal effects of investments, helping to determine whether neighboring regions' investments have spillover impacts on GDP per capita. By using these spatial econometric techniques, we can gain insights into regional economic dynamics and identify potential areas for targeted policy interventions.