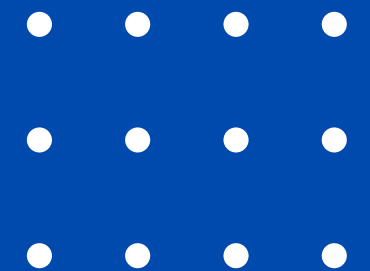


Proyecto **FINAL**

Andrea Hinojosa
Manuel Kanahuati
Andrés Robles



Nuestra pregunta

¿Qué variables socioeconómicas y ambientales inciden más en el PIB de cada país?



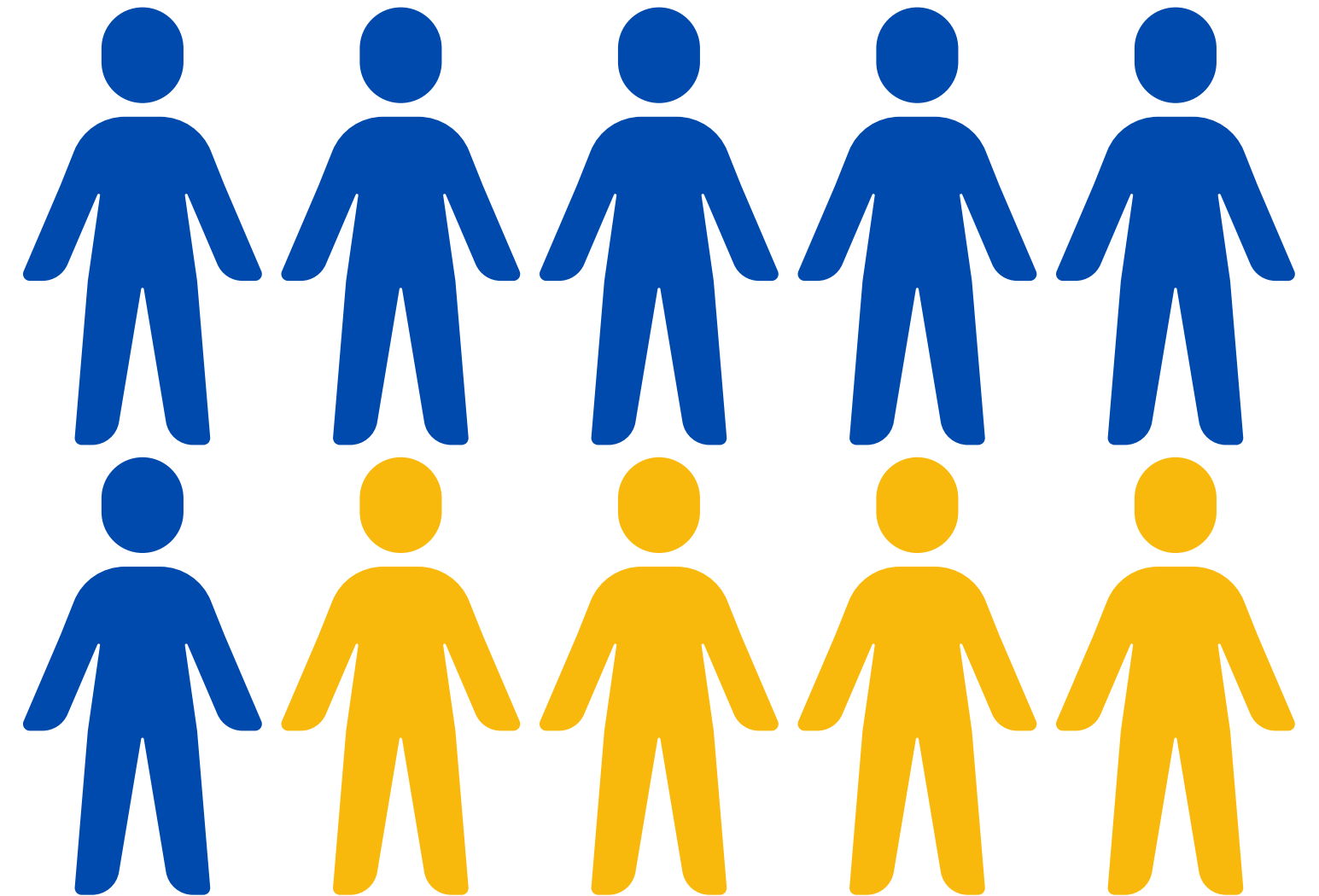
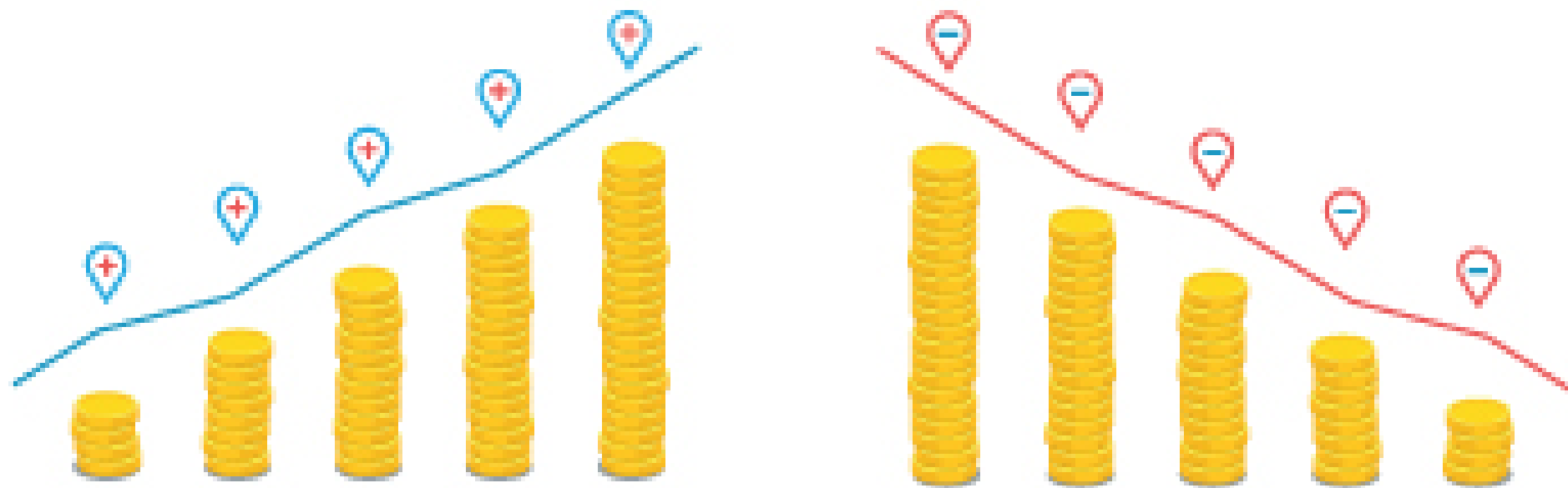
Background

- El PIB es la suma del valor (en dinero) de todos los bienes y servicios que genera un país durante un período.
- El cambio en el PIB a lo largo del tiempo es uno de los indicadores más importantes del crecimiento económico.
- Queremos saber qué actividades de un país tienen efecto en el valor del PIB.



Importancia

Es muy importante saber si la economía del país está creciendo o no, es decir, si se produjo más o menos que el año anterior. El cambio en el PIB a lo largo del tiempo es uno de los indicadores más importantes del crecimiento económico.



Fuentes de datos

Todos nuestros set de datos fueron obtenidos en el World Bank Data. Se extrajo de las 6 variables información de 74 países en el año 2011

PIB

Suma del valor (en dólares) de todos los bienes y servicios que genera un país.

Exportaciones

Valor de todos los bienes y servicios de mercado prestados al resto del mundo.

Impuestos

Recaudación impositiva: ansferencias obligatorias al Gobierno central con fines públicos

Market Cap

Suma del valor de las empresas públicas: precio de la acción por el número de acciones

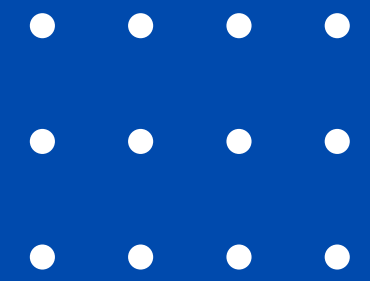
CO2

Emisiones en ton que genera un país de dióxido de carbono.

R&D

Gasto en investigación y desarrollo que hace un país.

Nuestras variables



Independientes

- Emisiones de CO2 por país
- Exportaciones hechas por país
- Impuestos recaudados por país
- Gasto en Investigación y desarrollo por país
- Valor de empresas públicas por país

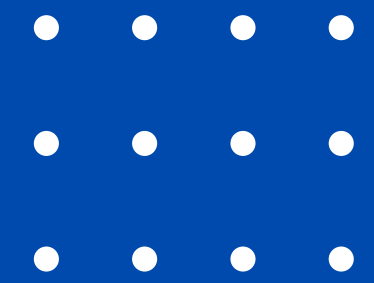
Dependiente

- PIB por país

Extra:

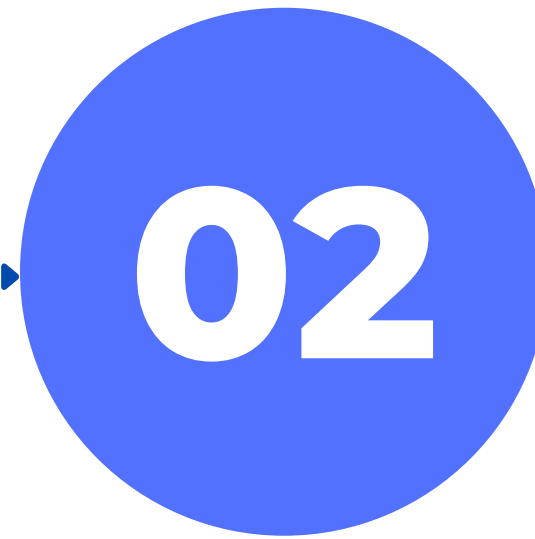
Población total por país para ayudarnos a escalar todas las variables a valor per capita

Preprocesamiento de datos



Elegir año

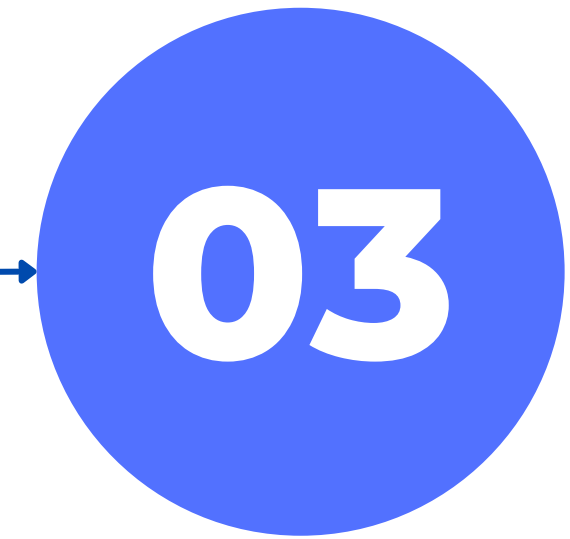
Se hizo una exploración de los datos para saber en qué año las 6 variables tenían menos valores nulos en el mayor número de países



Eliminar nulos

Se eliminaron las filas (países) que en el año 2011 tuvieron valores nulos en alguna de las 6 variables

Al final nos quedamos con 74 países



Mismas unidades

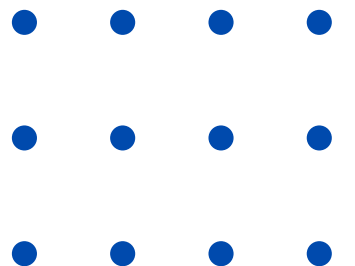
Cambiamos las unidades de todas las variables* para que todas estuvieran en US dólares per capita.

Aquí utilizamos la variable "extra" para escalar las variable a su valor per capita

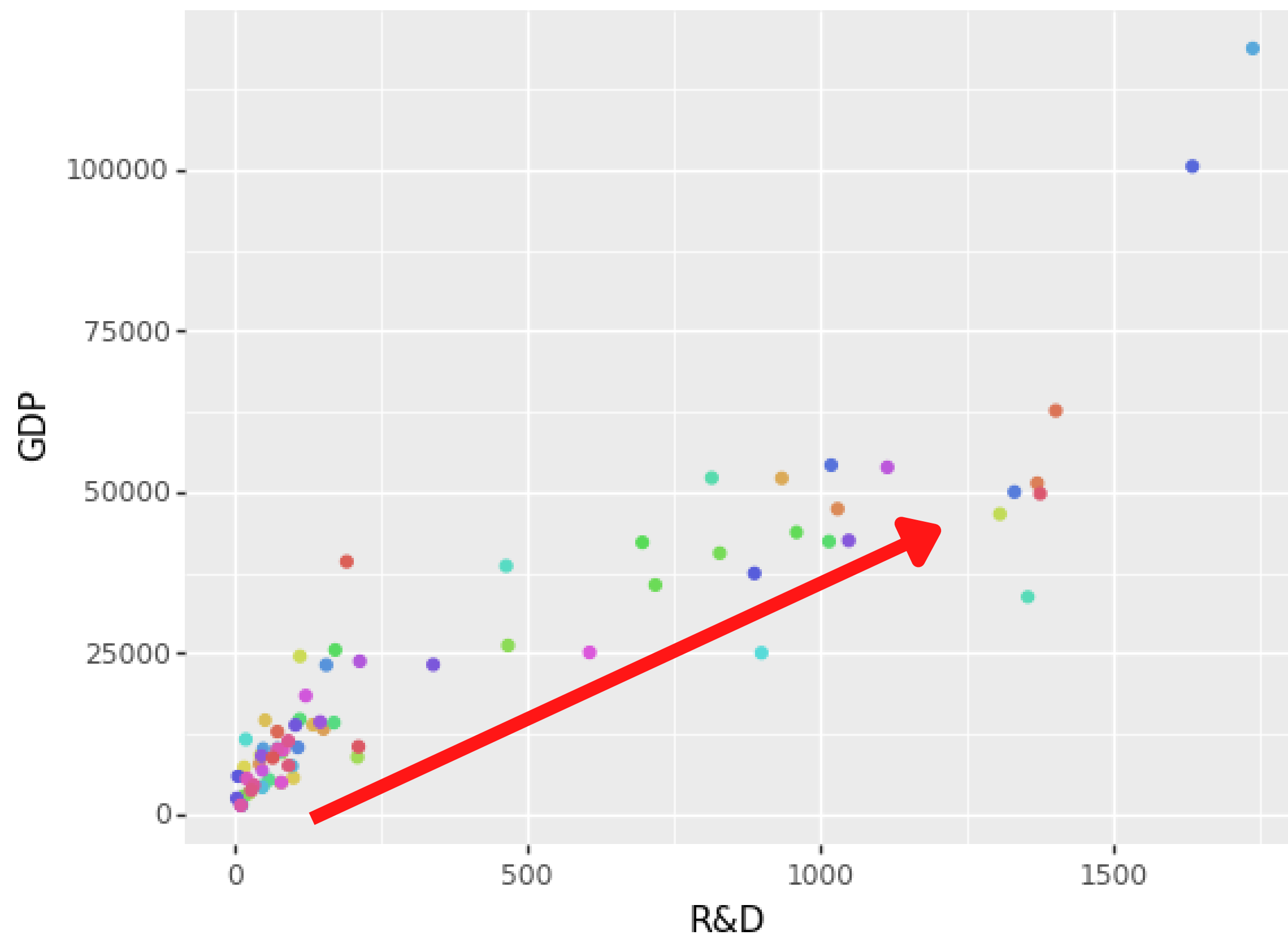
*menos CO2 esa la mantuvimos en tonCO2 per capita

Análisis exploratorio

Para analizar los datos que tenemos de cada variable independiente hicimos visualizaciones para observar la relación que cada una tiene con la variable dependiente (PIB)



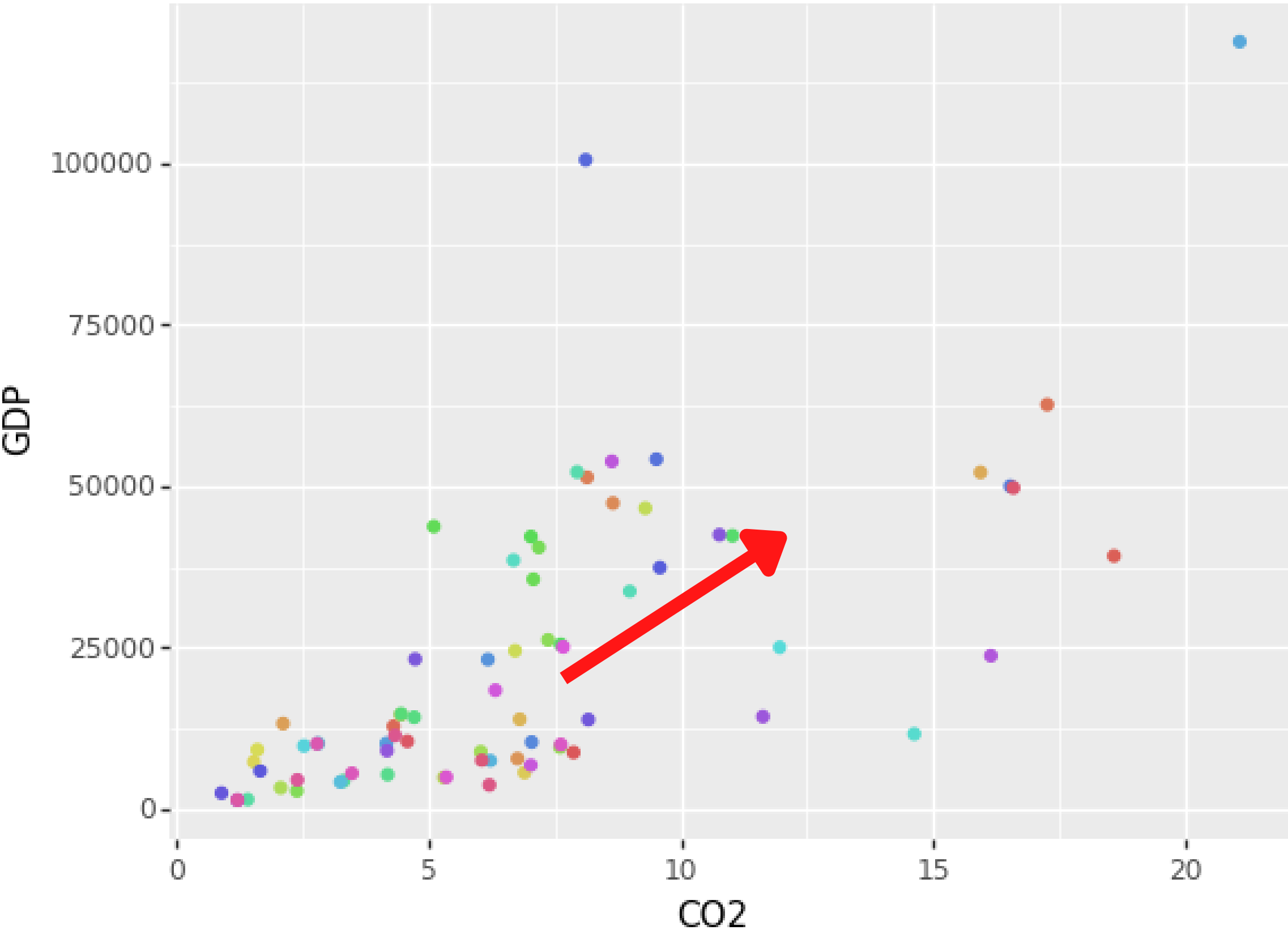
RELATION BETWEEN R&D AND GDP IN 2011



Country Code

ARE	ECS	LCN	SAU
ARG	EGY	LMY	SGP
AUS	EMU	LTE	SRB
AUT	EUU	LUX	SVK
BEL	FRA	MEX	SVN
BGR	GBR	MLT	TEA
BRA	GRC	MYS	TEC
CAN	HIC	NAC	THA
CEB	HRV	NLD	TLA
CHL	HUN	NOR	TSA
CHN	IBD	OED	TUN
COL	IBT	PER	TUR
CRI	IND	PHL	UKR
CYP	IRL	POL	UMC
DEU	ISR	PRT	USA
EAP	ITA	PST	WLD
EAR	KAZ	ROU	ZAF
EAS	KOR	RUS	
ECA	LAC	SAS	

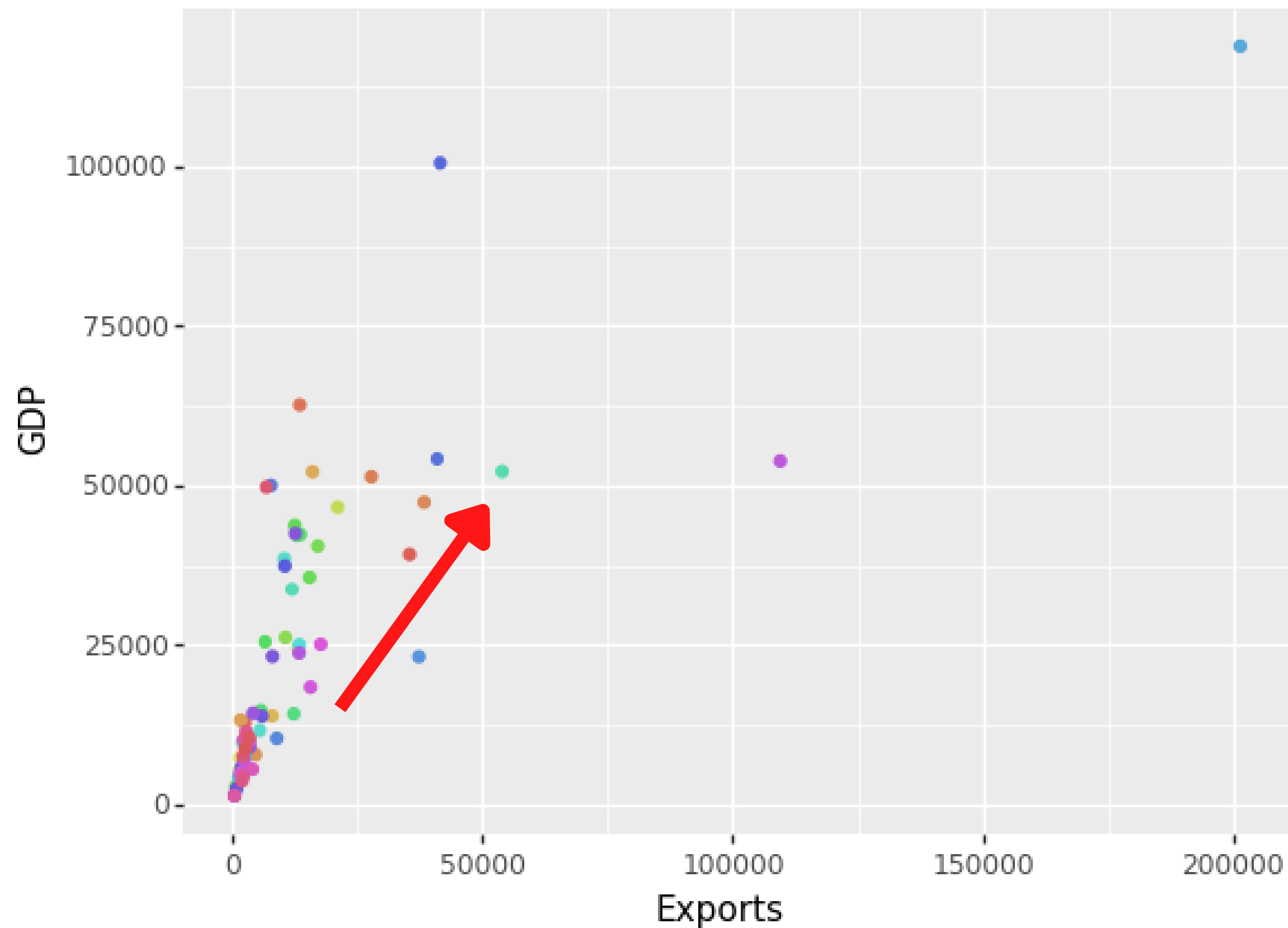
RELATION BETWEEN CO2 AND GDP IN 2011



Country Code

ARE	ECS	LCN	SAU
ARG	EGY	LMY	SGP
AUS	EMU	LTE	SRB
AUT	EUU	LUX	SVK
BEL	FRA	MEX	SVN
BGR	GBR	MLT	TEA
BRA	GRC	MYS	TEC
CAN	HIC	NAC	THA
CEB	HRV	NLD	TLA
CHL	HUN	NOR	TSA
CHN	IBD	OED	TUN
COL	IBT	PER	TUR
CRI	IND	PHL	UKR
CYP	IRL	POL	UMC
DEU	ISR	PRT	USA
EAP	ITA	PST	WLD
EAR	KAZ	ROU	ZAF
EAS	KOR	RUS	
ECA	LAC	SAS	

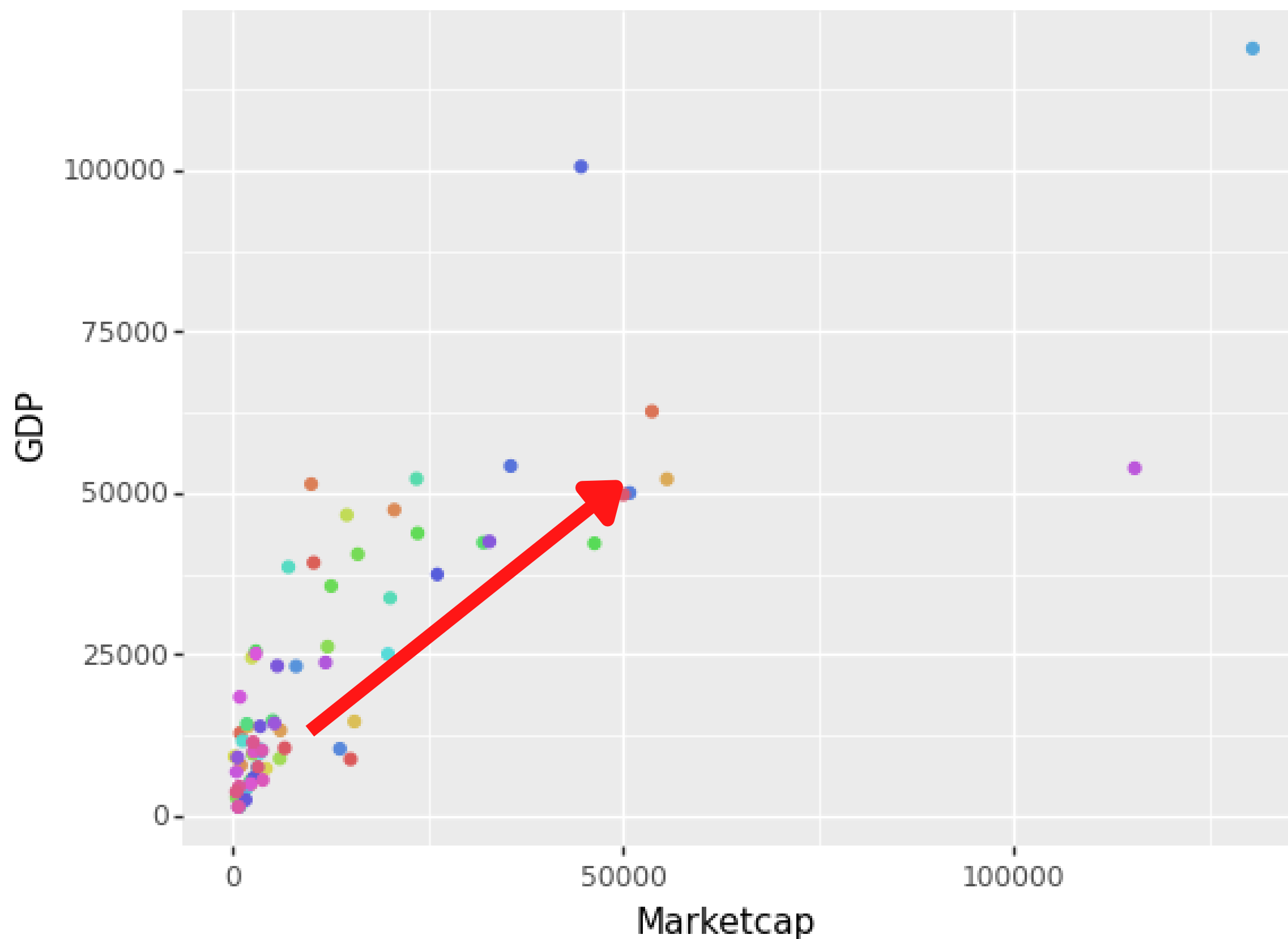
RELATION BETWEEN Exports AND GDP IN 2011



Country Code

ARE	ECS	LCN	SAU
ARG	EGY	LMY	SGP
AUS	EMU	LTE	SRB
AUT	EUU	LUX	SVK
BEL	FRA	MEX	SVN
BGR	GBR	MLT	TEA
BRA	GRC	MYS	TEC
CAN	HIC	NAC	THA
CEB	HRV	NLD	TLA
CHL	HUN	NOR	TSA
CHN	IBD	OED	TUN
COL	IBT	PER	TUR
CRI	IND	PHL	UKR
CYP	IRL	POL	UMC
DEU	ISR	PRT	USA
EAP	ITA	PST	WLD
EAR	KAZ	ROU	ZAF
EAS	KOR	RUS	
ECA	LAC	SAS	

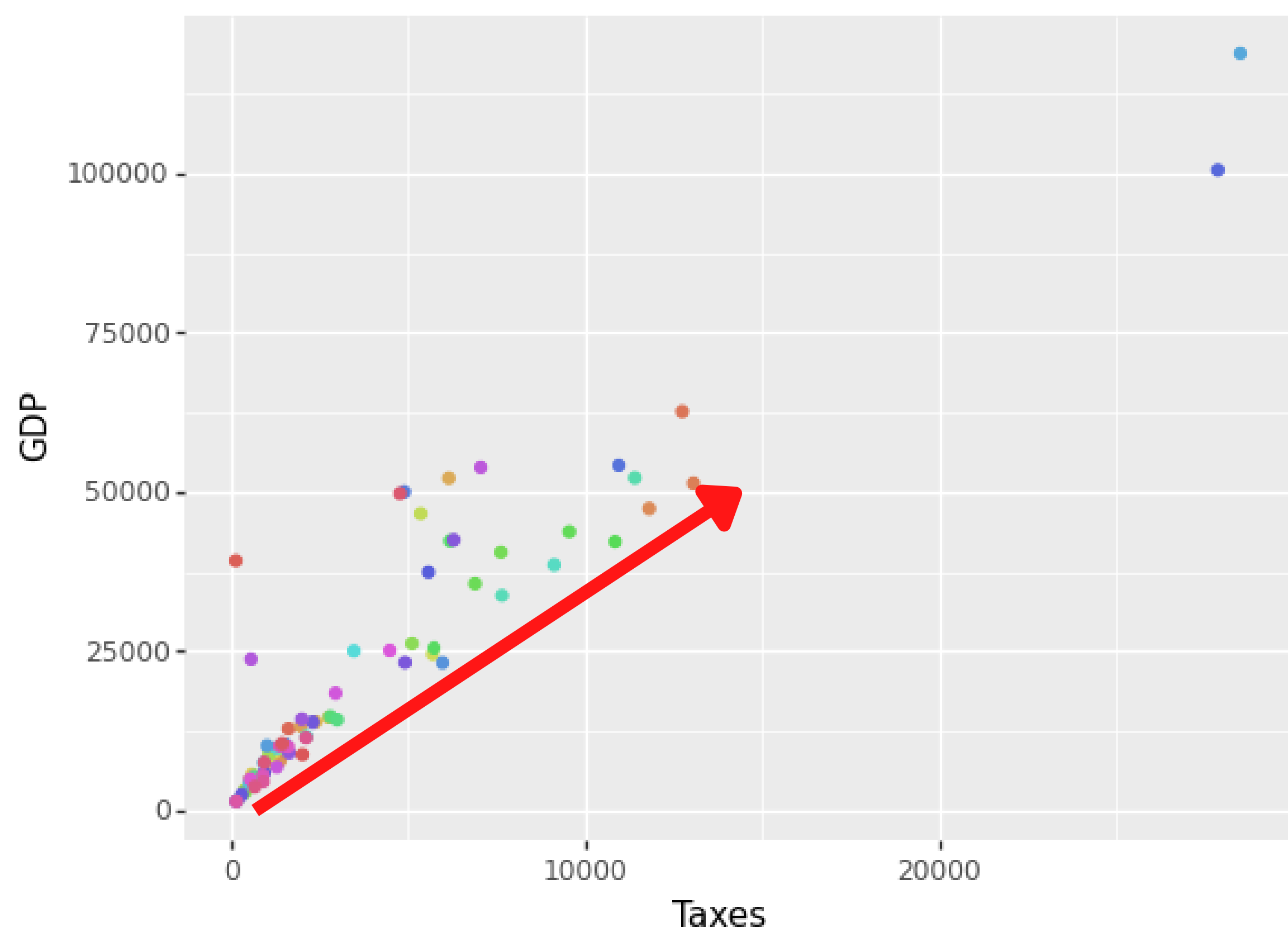
RELATION BETWEEN MARKETCAP AND GDP IN 2011



Country Code

ARE	ECS	LCN	SAU
ARG	EGY	LMY	SGP
AUS	EMU	LTE	SRB
AUT	EUU	LUX	SVK
BEL	FRA	MEX	SVN
BGR	GBR	MLT	TEA
BRA	GRC	MYS	TEC
CAN	HIC	NAC	THA
CEB	HRV	NLD	TLA
CHL	HUN	NOR	TSA
CHN	IBD	OED	TUN
COL	IBT	PER	TUR
CRI	IND	PHL	UKR
CYP	IRL	POL	UMC
DEU	ISR	PRT	USA
EAP	ITA	PST	WLD
EAR	KAZ	ROU	ZAF
EAS	KOR	RUS	
ECA	LAC	SAS	

RELATION BETWEEN TAX COLLECTION AND GDP IN 2011



Country Code			
ARE	ECS	LCN	SAU
ARG	EGY	LMY	SGP
AUS	EMU	LTE	SRB
AUT	EUU	LUX	SVK
BEL	FRA	MEX	SVN
BGR	GBR	MLT	TEA
BRA	GRC	MYS	TEC
CAN	HIC	NAC	THA
CEB	HRV	NLD	TLA
CHL	HUN	NOR	TSA
CHN	IBD	OED	TUN
COL	IBT	PER	TUR
CRI	IND	PHL	UKR
CYP	IRL	POL	UMC
DEU	ISR	PRT	USA
EAP	ITA	PST	WLD
EAR	KAZ	ROU	ZAF
EAS	KOR	RUS	
ECA	LAC	SAS	

MODELO: REGRESIÓN LINEAL

- Es un **modelo supervisado** al trabajar con datos etiquetados y lo que queremos es encontrar una función, que con nuestras variables independientes, le asignen al PIB un valor de salida (en este caso número).
- Es un **problema de regresión** porque lo que buscamos es un número por cada variable independiente que nos ayude a explicar el comportamiento del PIB.
- **Inferencia** ya que con los coeficientes que nos de la regresión como resultado podremos saber cuáles de nuestras variables independientes tienen un efecto más grande en el PIB.

MODELO: REGRESIÓN LINEAL

- División del data set:
 - 80% de los datos se utilizaron para el entrenamiento
 - 20% de los datos se utilizaron para las pruebas

Intercepto

1585.24

Coeficientes

R&D Expense: 1.26677200e+01

Exports: 6.65679063e-03

MARKETCAP: 1.28116971e-01

Tax collection: 2.34192587e+00

CO2 emissions: 6.52482095e+02

Score r2

0.8467

INTERPRETACIÓN RESULTADOS

- Coeficientes

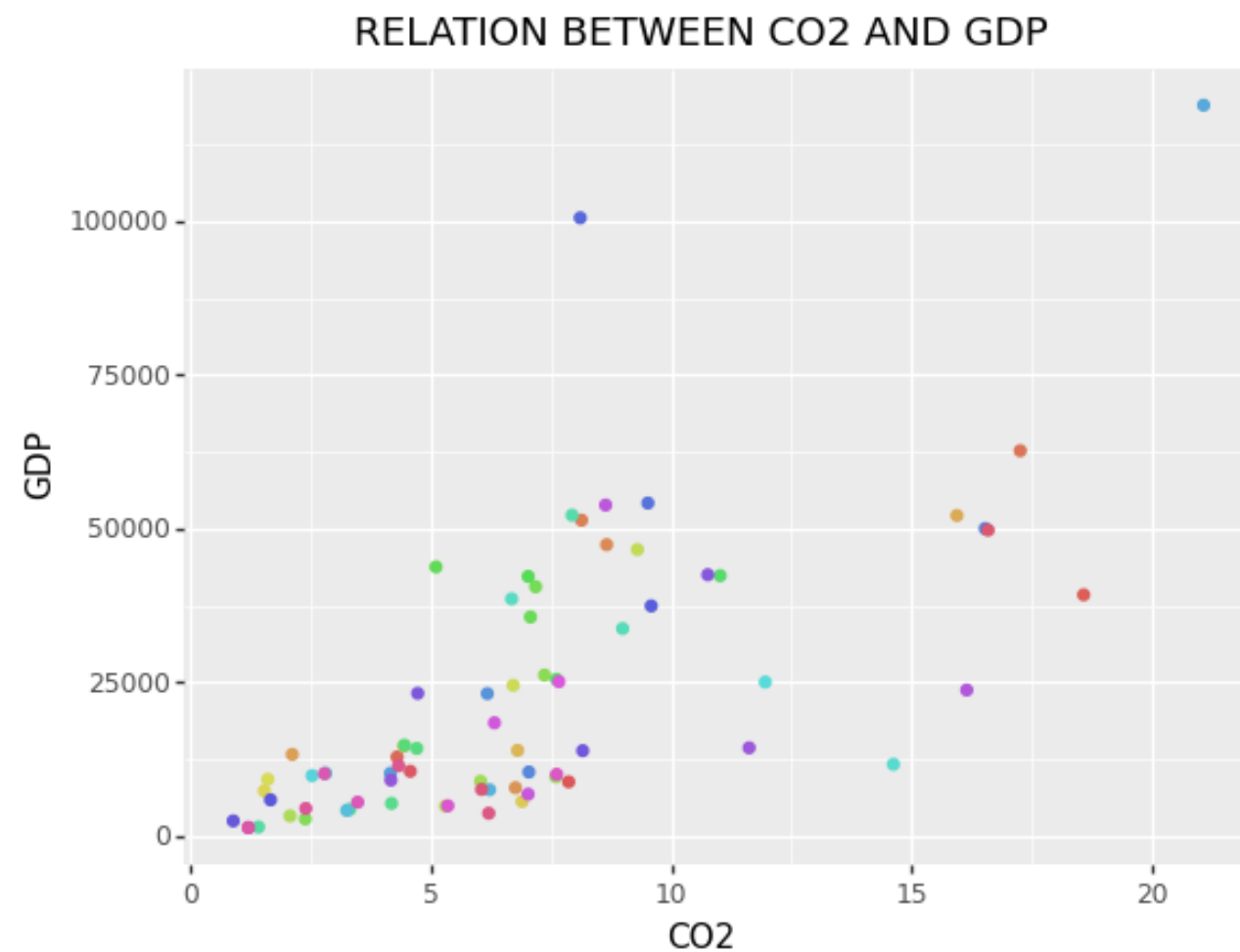
- Gasto R&D: Si el gasto en R&D aumenta 1 dólar per capita, el PIB incrementará $1.26677200e+01$ en dólares per capita
- Exports: Si el valor de las exportaciones aumenta 1 dólar per capita, el PIB incrementará en $6.65679063e-03$ dólares per capita
- MARKETCAP: Si el valor de MARKETCAP aumenta 1 dólar per capita, el PIB incrementará en $1.28116971e-01$ dólares per capita
- Tax collection: Si la recolección de impuestos aumenta 1 dólar per capita, el PIB incrementará en $2.34192587e+00$ dólares per capita
- CO2 emissions: Si las emisiones de CO2 aumenta en 1 tonCO2 per capita, el PIB incrementará en $6.52482095e+02$ dólares per capita

INTERPRETACIÓN RESULTADOS

- Intercept: si todas nuestras variables independientes son igual a carp, el valor promedio del PIB sería 1585 USD per capita
- R^2 score = 0.8467
 - El mejor resultado posible es 1
 - Califica qué tan bueno es nuestro modelo para predecir

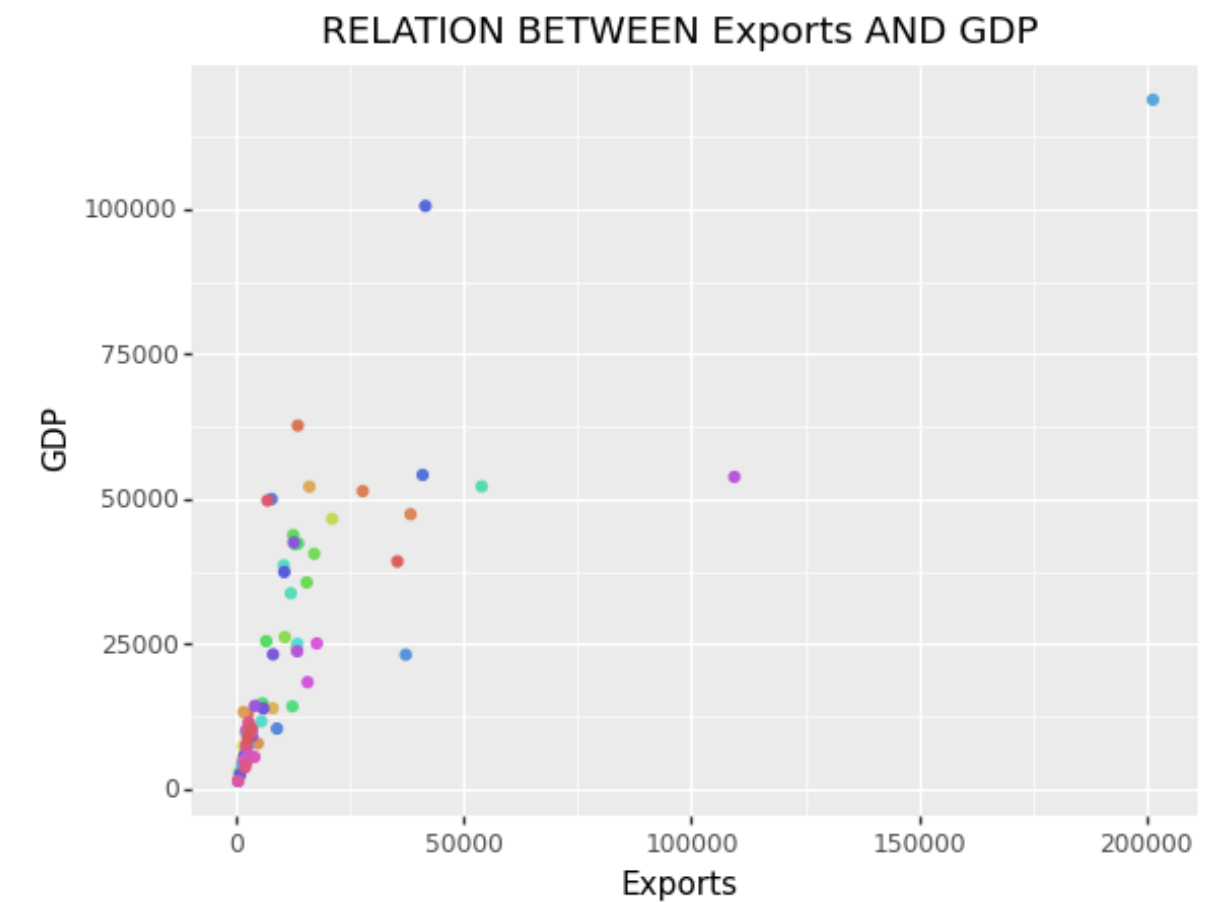
CONCLUSIONES

- La variable dependiente que más relación tiene con el PIB es el CO2 (emisiones en toneladas de CO2 per capita)



CO2 emissions: 6.52482095e+02

- La variable dependiente que menos relación tiene con el PIB son las exportaciones (USD per capita)



Exports: 6.65679063e-03

GRACIAS

