

Trabajo-Final-Grupo8

Roberto Cardozo, Ricardo Sanchez, Sebastian Hinostroza, Soledad Lopez

2024-04-29

Analisis de Datos de Producción de Yacimientos no convencionales en Argentina

Introduccion

Argentina ha experimentado un crecimiento gradual en la producción de yacimientos no convencionales. Esto se debe principalmente a que, en sus inicios, el país carecía de la tecnología y el conocimiento necesarios para explotar estos recursos de manera eficiente y rentable. La inversión inicial requerida era alta, con perspectivas de ganancias inciertas. Estados Unidos, como pionero en la explotación de YNC, sirvió como modelo para que otros países, como Argentina, pudieran alcanzar el éxito en este ámbito.

En este informe, analizaremos el crecimiento de la producción de YNC en Argentina a lo largo de los años. Para ello, utilizaremos un conjunto de datos (dataset) que detalla la información de todos los pozos perforados desde el año 2016 hasta la actualidad. La gran cantidad de datos disponible nos exige ser sumamente meticulosos en la selección de la información a analizar.

Datasets importado de la Secretaria de energia

```
library(pacman) # Gestor de paquetes en R
p_load("readr") # Biblioteca para importar archivos csv
p_load("dplyr") # Biblioteca para manipular datos
p_load("ggplot2") # Biblioteca para graficar datos
```

```
library(readr)
produccin_de_pozos_de_gas_y_petrleo_no_convencional <- read_csv("C:/Users/gladly/Downloads/produccin-de-
  col_types = cols(coordenadax = col_number(),
    coordenaday = col_number()))
```

filtrado de data

Teniendo en cuenta que el datasets está compuesto por más de 295.306 registros de pozos, con el fin de analizar la información relevante, procederemos a filtrar los datos en función de las características más importantes. El primer filtro será eliminar los datos del año 2024, ya que al ser un año en curso, los datos disponibles podrían no ser representativos de un año completo y generar una distorsión en la tendencia de producción anual acumulada al compararla con años anteriores.

```
filtrardata <- subset(produccion_de_pozos_de_gas_y_petrleo_no_convencional, !anio %in% 2024)
filtrardata [1:8,1:7]
```

```
## # A tibble: 8 x 7
##   idempresa anio   mes idpozo prod_pet prod_gas prod_agua
##   <chr>     <dbl> <dbl> <dbl>   <dbl>   <dbl>   <dbl>
## 1 YSUR      2016     1 129863     0     172.    153.
## 2 YSUR      2017     1 131430    28.6    648.     20.0
## 3 YSUR      2016     1 130515     0      20.4     17.8
## 4 YSUR      2017     1 152786    24.8    512.     20.6
## 5 YSUR      2016     1 137599    52.8    548.     38.5
## 6 YSUR      2017     1 135208     0     105.      0
## 7 YSUR      2016     1 143845    47.6    551.     64.4
## 8 YSUR      2017     1 131345     5.88   306.      5.98
```

Produccion de Petroleo de yacimientos no convencional a lo largo de los años.

En Argentina, la explotación de yacimientos no convencionales ha crecido de manera progresiva en los últimos años, ya que no se disponía de los conocimientos y la tecnología necesaria para poder obtener un rendimiento económico a una inversión de tal magnitud. Esto se refleja en el siguiente gráfico:

```
grupos_anios <- group_by(filtrardata,anio)
grupos_anios[1:6,1:6]
```

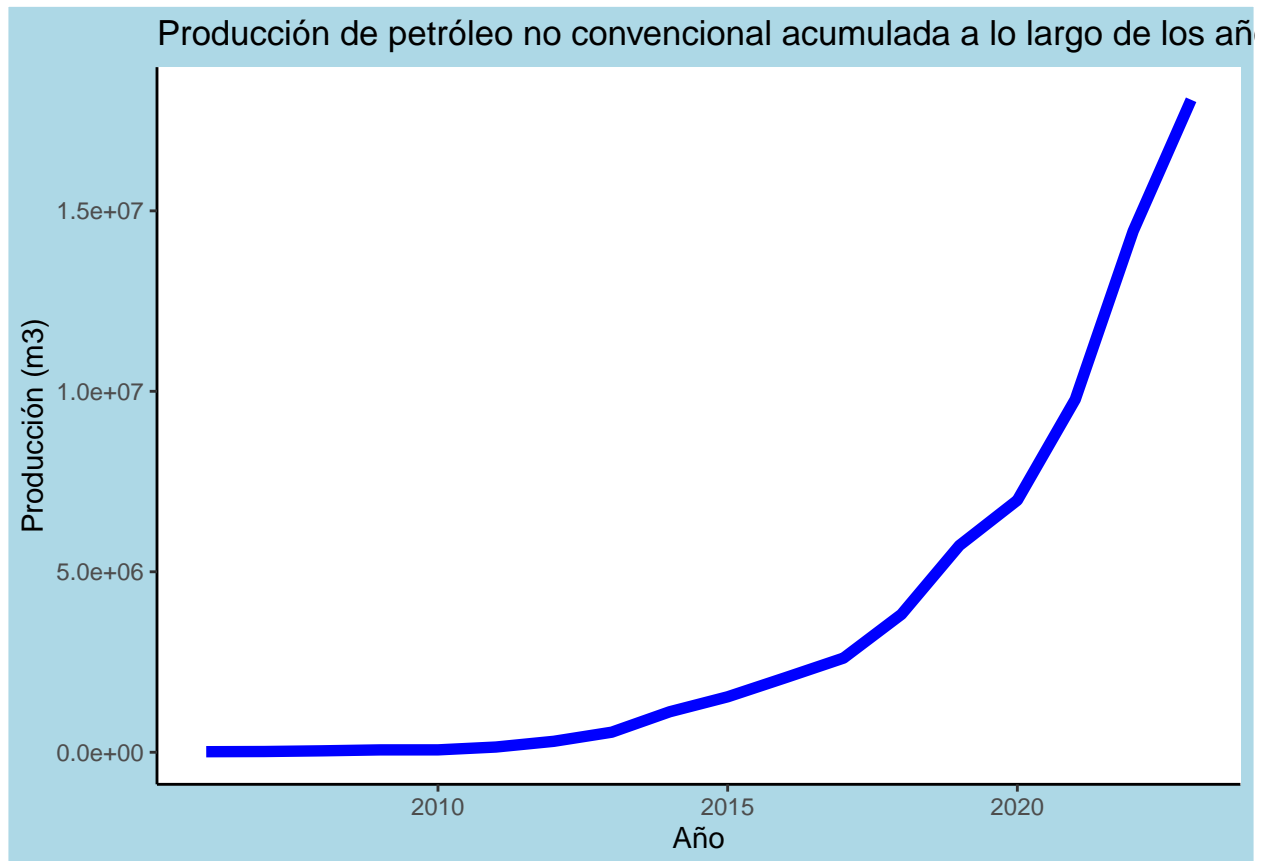
```
## # A tibble: 6 x 6
## # Groups:   anio [2]
##   idempresa anio   mes idpozo prod_pet prod_gas
##   <chr>     <dbl> <dbl> <dbl>   <dbl>   <dbl>
## 1 YSUR      2016     1 129863     0     172.
## 2 YSUR      2017     1 131430    28.6    648.
## 3 YSUR      2016     1 130515     0      20.4
## 4 YSUR      2017     1 152786    24.8    512.
## 5 YSUR      2016     1 137599    52.8    548.
## 6 YSUR      2017     1 135208     0     105.
```

```
promedioanios<-summarise(grupos_anios,mean = mean(prod_pet), sum = sum(prod_pet), n = n())
promedioanios[11:18,]
```

```
## # A tibble: 8 x 4
##   anio mean      sum      n
##   <dbl> <dbl>   <dbl> <int>
## 1 2016 120. 2068063. 17225
## 2 2017 126. 2609358. 20725
## 3 2018 154. 3821723. 24883
## 4 2019 198. 5728448. 28905
## 5 2020 220. 6981549. 31704
## 6 2021 283. 9776037. 34494
## 7 2022 371. 14439821. 38906
## 8 2023 419. 18077598. 43195
```

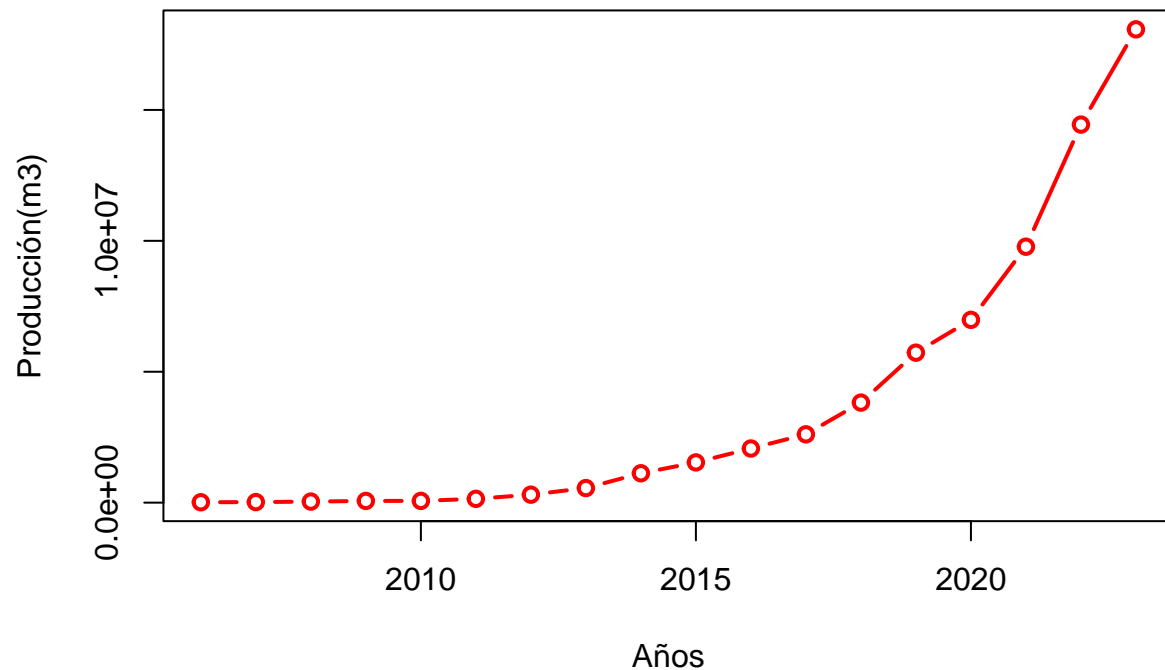
```
ggplot(promedioanios, aes(x = anio, y = sum)) +
  geom_line(size = 2, color = "blue") +
  labs(title = "Producción de petróleo no convencional acumulada a lo largo de los años", x = "Año", y = "Producción (m3)") +
  theme_classic() +
  theme(plot.background = element_rect(fill = "lightblue"))
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```



```
plot (promedioanios$anio,promedioanios$sum, main= "Produccion de Petroleo no convencional a lo largo de los años")
```

Produccion de Petroleo no convencional a lo largo de los años



Provincias productoras

```
grupo_provincias<-group_by(filtrardata,provincia)
grupo_provincias[1:6,1:8]
```

```
## # A tibble: 6 x 8
##   idempresa  anio    mes idpozo prod_pet prod_gas prod_agua iny_agua
##   <chr>      <dbl> <dbl> <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
## 1 YSUR      2016     1 129863      0      172.     153.      0
## 2 YSUR      2017     1 131430     28.6     648.     20.0      0
## 3 YSUR      2016     1 130515      0      20.4     17.8      0
## 4 YSUR      2017     1 152786     24.8     512.     20.6      0
## 5 YSUR      2016     1 137599     52.8     548.     38.5      0
## 6 YSUR      2017     1 135208      0     105.      0      0
```

```
promedio_prov<-summarise(grupo_provincias,Produccion_promedio = mean(prod_pet),produccion_total= sum(pr
ordenadop <- promedio_prov %>%
  arrange(desc(produccion_total))
ordenadop
```

```
## # A tibble: 6 x 4
##   provincia Produccion_promedio produccion_total    n
##   <chr>      <dbl>          <dbl>    <int>
```

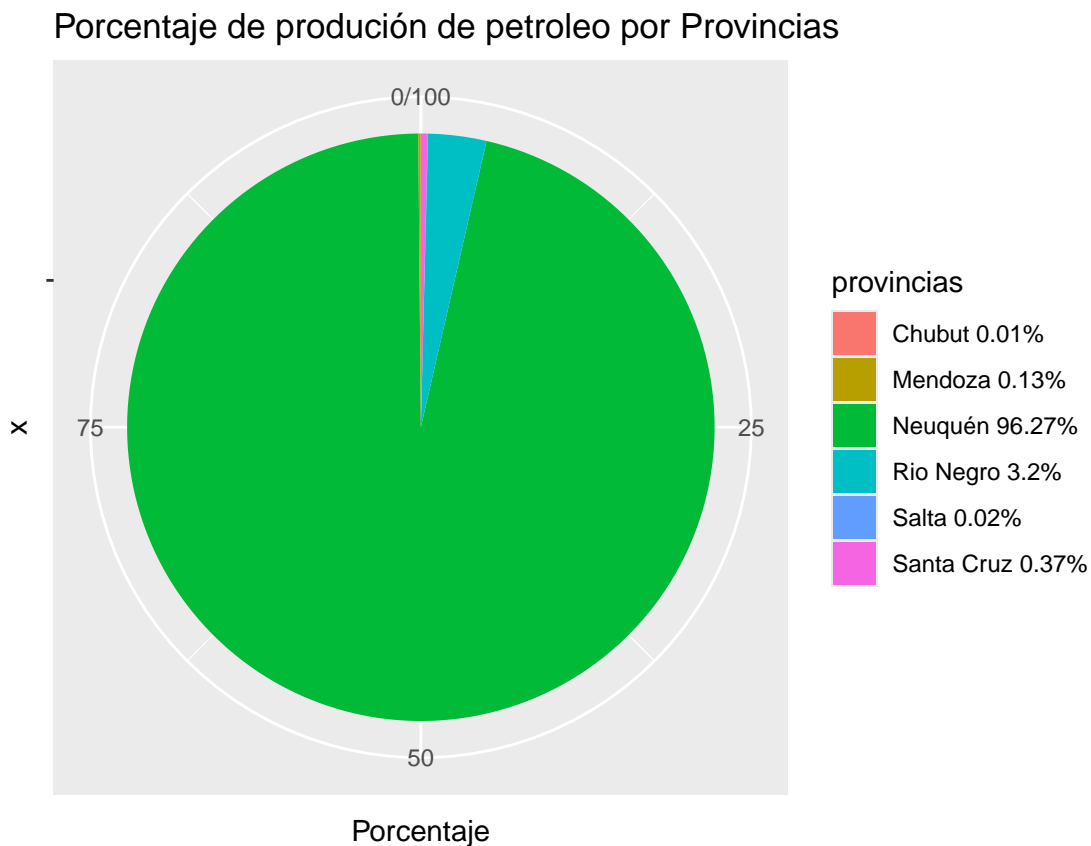
## 1 Neuquén	260.	64864490.	249250
## 2 Rio Negro	96.9	2155076.	22248
## 3 Santa Cruz	25.3	252313.	9975
## 4 Mendoza	42.0	90347.	2149
## 5 Salta	83.0	10627.	128
## 6 Chubut	26.6	4099.	154

```

proporciones <- ordenadop$produccion_total # creamos un vector con proporciones
provincias <- ordenadop$provincia # vector con etiquetas
Porcentaje <- round(proporciones/sum(proporciones)*100,2)
provincias <- paste(provincias, Porcentaje) # Añadimos porcentajes a etiquetas
provincias <- paste(provincias,"%",sep="")

ggplot(ordenadop , aes(x="", y= Porcentaje , fill=provincias)) +
  geom_bar(stat = "identity")+
  coord_polar("y", start=0) +
  labs(title = "Porcentaje de producción de petroleo por Provincias ")

```



Las Formaciones mas importantes

Uno de los aspectos más importantes a analizar son las formaciones. En la actualidad, la formación más importante de Argentina es Vaca Muerta, que produce más del 90% del petróleo no convencional, seguida por las formaciones Lajas y Punta Rosada con un porcentaje mucho menor. Para determinar esto a partir del conjunto de datos, utilizamos la función `groupby()`, para agrupar todos los datos por formacio, y la función `summarise()`, para calcular la producción promedio por formación y la producción acumulada.

```
grupo_Formacion<-group_by(filtrardata,formacion)
grupo_Formacion[1:8,25:29]
```

```
## # A tibble: 8 x 5
## # Groups:   formacion [2]
##   profundidad formacion idareapermisiconcesion areapermisiconcesion
##   <dbl> <chr> <chr> <chr>
## 1 3500 precuyo ANC ANTICLINAL CAMPAMENTO
## 2 3800 lajas FEO ESTACION FERNANDEZ ORO
## 3 3428 precuyo ANC ANTICLINAL CAMPAMENTO
## 4 2940 precuyo SDD AL SUR DE LA DORSAL
## 5 3764 lajas FEO ESTACION FERNANDEZ ORO
## 6 3166 precuyo ANC ANTICLINAL CAMPAMENTO
## 7 3783 lajas FEO ESTACION FERNANDEZ ORO
## 8 3806 lajas FEO ESTACION FERNANDEZ ORO
## # i 1 more variable: idareayacimiento <chr>
```

Promedio por formación

```
promedioform<-summarise(grupo_Formacion,Produccion_promedio = mean(prod_pet),produccion_total= sum(prod_pet))
ordenado <- promedioform %>%
  arrange(desc(produccion_total))
ordenado[1:10,]
```

```
## # A tibble: 10 x 4
##   formacion Produccion_promedio produccion_total n
##   <chr> <dbl> <dbl> <int>
## 1 vaca muerta 481. 63145983. 131249
## 2 lajas 39.9 2355935. 59090
## 3 punta rosada 98.2 623470. 6348
## 4 mulichinco 14.7 393922. 26720
## 5 magallanes 24.7 235622. 9541
## 6 quintuco 19.8 168479. 8509
## 7 los molles 7.40 157675. 21321
## 8 precuyo 12.1 146394. 12092
## 9 agrio 20.1 68190. 3398
## 10 lotena 14.9 19202. 1285
```

```
porcentaje<- reframe(ordenado,formacion=ordenado$formacion, Porcentaje_Producción=round((ordenado$produccion_total/ordenado$Produccion_promedio)*100))
porcentaje [1:10,]
```

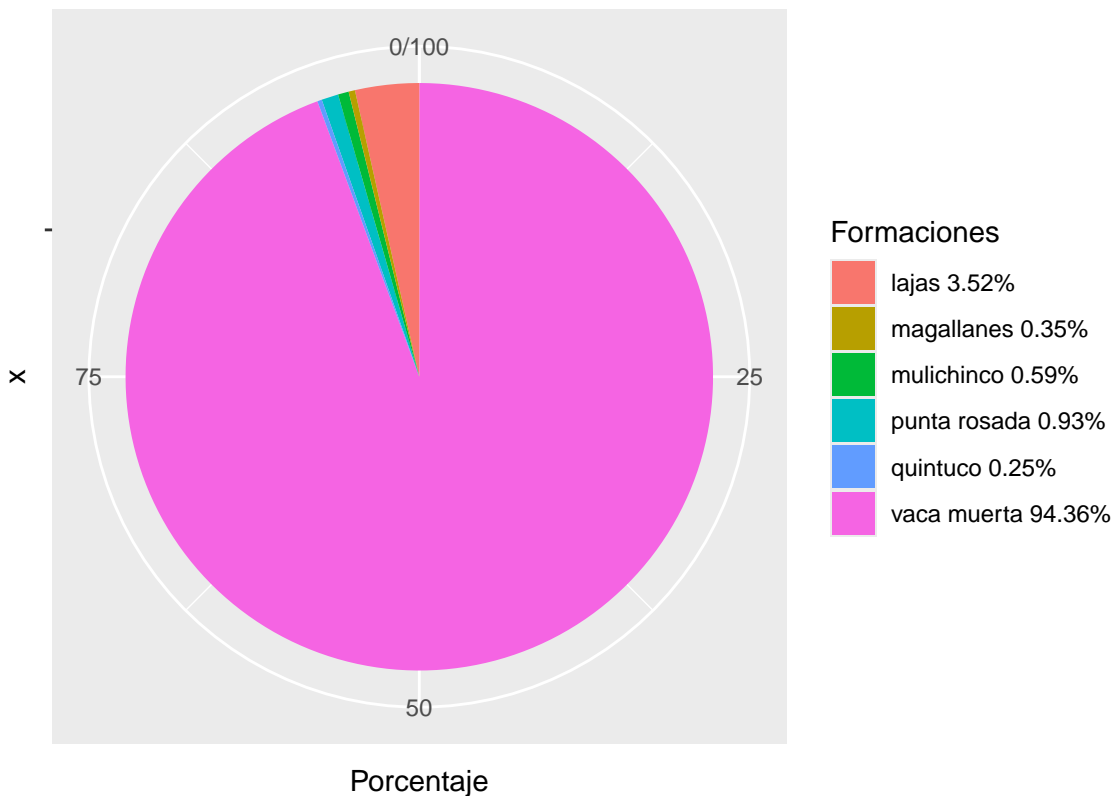
```
## # A tibble: 10 x 2
##   formacion Porcentaje_Producción
##   <chr> <dbl>
## 1 vaca muerta 93.7
## 2 lajas 3.50
## 3 punta rosada 0.925
## 4 mulichinco 0.585
## 5 magallanes 0.35
## 6 quintuco 0.25
```

```
## 7 los molles          0.234
## 8 precuyo             0.217
## 9 agrio                0.101
## 10 lotena             0.028
```

```
formacionesimp <- head(ordenado,6) # con la funcion head tomamos solo las 6 formaciones mas importantes
proporciones <- formacionesimp$produccion_total # creamos un vector con proporciones
Formaciones <- formacionesimp$formacion # vector con etiquetas
Porcentaje <- round(proporciones/sum(proporciones)*100,2)
Formaciones <- paste(Formaciones, Porcentaje) # Añadimos porcentajes a etiquetas
Formaciones <- paste(Formaciones,"%",sep="")

ggplot(head(ordenado,6), aes(x="", y= Porcentaje , fill=Formaciones)) +
  geom_bar(stat = "identity")+
  coord_polar("y", start=0) +
  labs(title = "Porcentaje de producción de petroleo por Formacion")
```

Porcentaje de producción de petroleo por Formacion



Para mayor claridad representamos solo las 3 formaciones mas importantes en el siguiente grafico de torta

```
formacionesimp <- head(ordenado,3)
proporciones <- formacionesimp$produccion_total # creamos un vector con proporciones
etiquetas <- formacionesimp$formacion # vector con etiquetas
pct <- round(proporciones/sum(proporciones)*100)
etiquetas <- paste(etiquetas, pct) # Añadimos porcentajes a etiquetas
etiquetas <- paste(etiquetas,"%",sep="")
leyenda <- formacionesimp$formacion # Añadimos el símbolo de %
```

```
pie(proporciones, labels = leyenda,
    col=rainbow(length(leyenda)),
    main="Formaciones mas importantes ", radius = 1)

# Añadimos un cuadro con leyendas
legend("topright", etiquetas, cex = 1,
      fill = rainbow(length(proporciones)))
```



Producción de Petroleo en pandemia

Al contar con un Datasets con tanta información a lo largo de los años, otro aspecto que nos resulta de suma importancia analizar es la caída de la producción de petróleo durante la pandemia del año 2020, la cual tuvo un efecto significativo en la producción de gas natural, convirtiéndolo en una alternativa más atractiva para la generación de energía y la calefacción. El gas natural a menudo se considera un combustible más limpio en comparación con el carbón, y su menor precio lo convirtió en una opción más económica para muchos consumidores.

Para poder contrastar estas dos situaciones, filtraremos el conjunto de datos en tres años: 2019, 2020 y 2023. Elegimos un año anterior y posterior a la pandemia para poder visualizar la diferencia entre años donde el mercado se comportó de manera normal y un año con una situación global compleja.

```
library(dplyr)

año2019 <- filtrardata%>%
```



```

filter(anio == 2019)
meses<-group_by(año2019,mes)
produccion_mensual<-summarise(meses,Promedio_mensual = mean(prod_pet),produccion_total= sum(prod_pet),n
produccion_mensual

```

```

## # A tibble: 12 x 4
##   mes Promedio_mensual produccion_total    n
##   <dbl>          <dbl>          <dbl> <int>
## 1     1           196.          421050.  2143
## 2     2           166.          381155.  2295
## 3     3           174.          398494.  2290
## 4     4           183.          429853.  2350
## 5     5           193.          455964.  2362
## 6     6           180.          430696.  2392
## 7     7           194.          469966.  2425
## 8     8           220.          540583.  2462
## 9     9           210.          527904.  2513
## 10    10           215.          545731.  2539
## 11    11           211.          538311.  2556
## 12    12           228.          588742.  2578

```

```

año2020<- filtrardata%>%
  filter(anio == 2020)
meses<-group_by(año2020,mes)
produccion_mensual2020<-summarise(meses,Promedio_mensual = mean(prod_pet),produccion_total= sum(prod_pe
produccion_mensual2020

```

```

## # A tibble: 12 x 4
##   mes Promedio_mensual produccion_total    n
##   <dbl>          <dbl>          <dbl> <int>
## 1     1           228.          593488.  2598
## 2     2           222.          581147.  2614
## 3     3           245.          641548.  2619
## 4     4           180.          471648.  2619
## 5     5           188.          492849.  2621
## 6     6           224.          588446.  2628
## 7     7           230.          607333.  2646
## 8     8           231.          611160.  2649
## 9     9           211.          561005.  2661
## 10    10           225.          599847.  2667
## 11    11           221.          594029.  2683
## 12    12           237.          639048.  2699

```

```

año2023<- filtrardata%>%
  filter(anio == 2023)
meses<-group_by(año2023,mes)
produccion_mensual23<-summarise(meses,Promedio_mensual = mean(prod_pet),produccion_total= sum(prod_pet)
produccion_mensual23

```

```

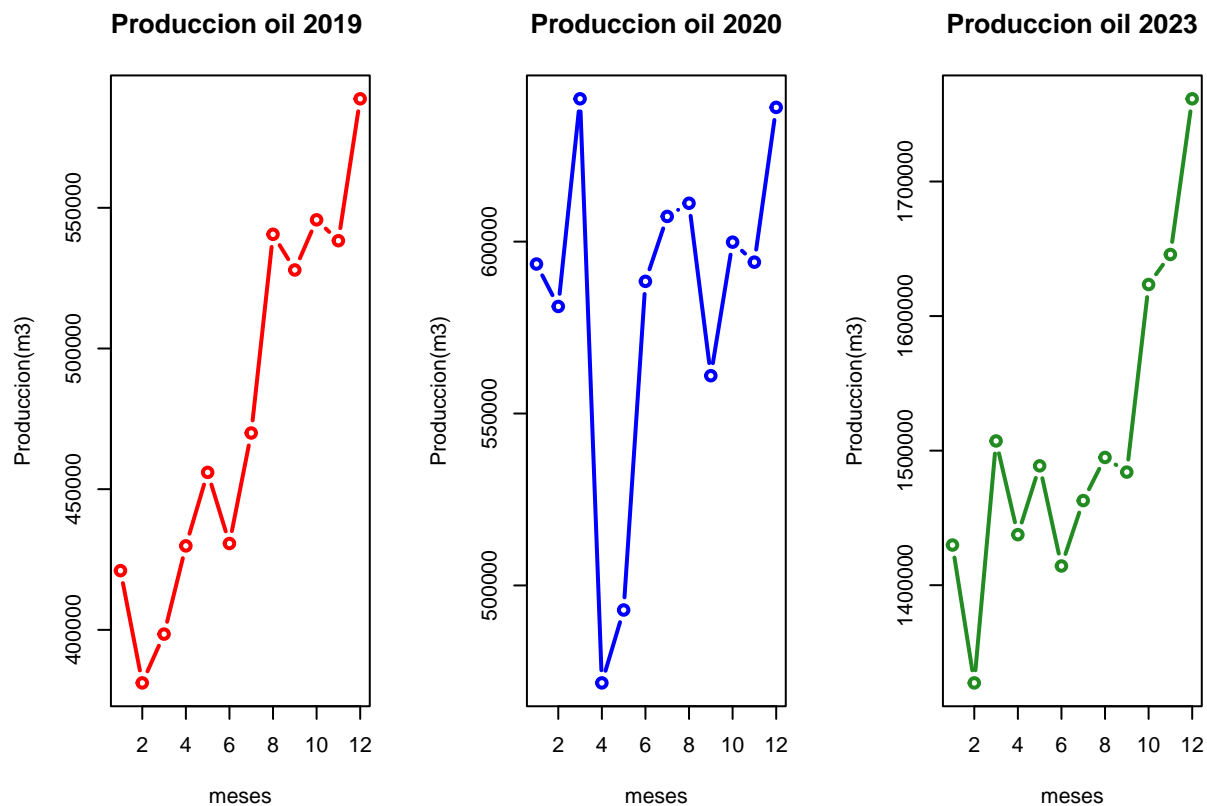
## # A tibble: 12 x 4
##   mes Promedio_mensual produccion_total    n
##   <dbl>          <dbl>          <dbl> <int>

```

##	1	1	416.	1429855.	3436
##	2	2	383.	1327399.	3463
##	3	3	433.	1507197.	3484
##	4	4	410.	1437563.	3510
##	5	5	421.	1488712.	3539
##	6	6	395.	1414278.	3583
##	7	7	406.	1462892.	3605
##	8	8	410.	1494930.	3644
##	9	9	403.	1484057.	3681
##	10	10	438.	1623454.	3706
##	11	11	439.	1645782.	3746
##	12	12	464.	1761479.	3798

Producción de Petroleo durante pandemia

```
par(mfrow=c(1,3))
plot (produccion_mensual$mes,produccion_mensual$produccion_total, main= "Produccion oil 2019 ",xlab="meses",col="red")
plot (produccion_mensual2020$mes,produccion_mensual2020$produccion_total ,main="Produccion oil 2020",xlab="meses",col="blue")
plot (produccion_mensual23$mes,produccion_mensual23$produccion_total,main="Produccion oil 2023",xlab="meses",col="green")
```



El gráfico muestra una clara disminución de la producción de petróleo en abril, coincidiendo con el pico de la pandemia

Produccion de Gas durante pandemia

```

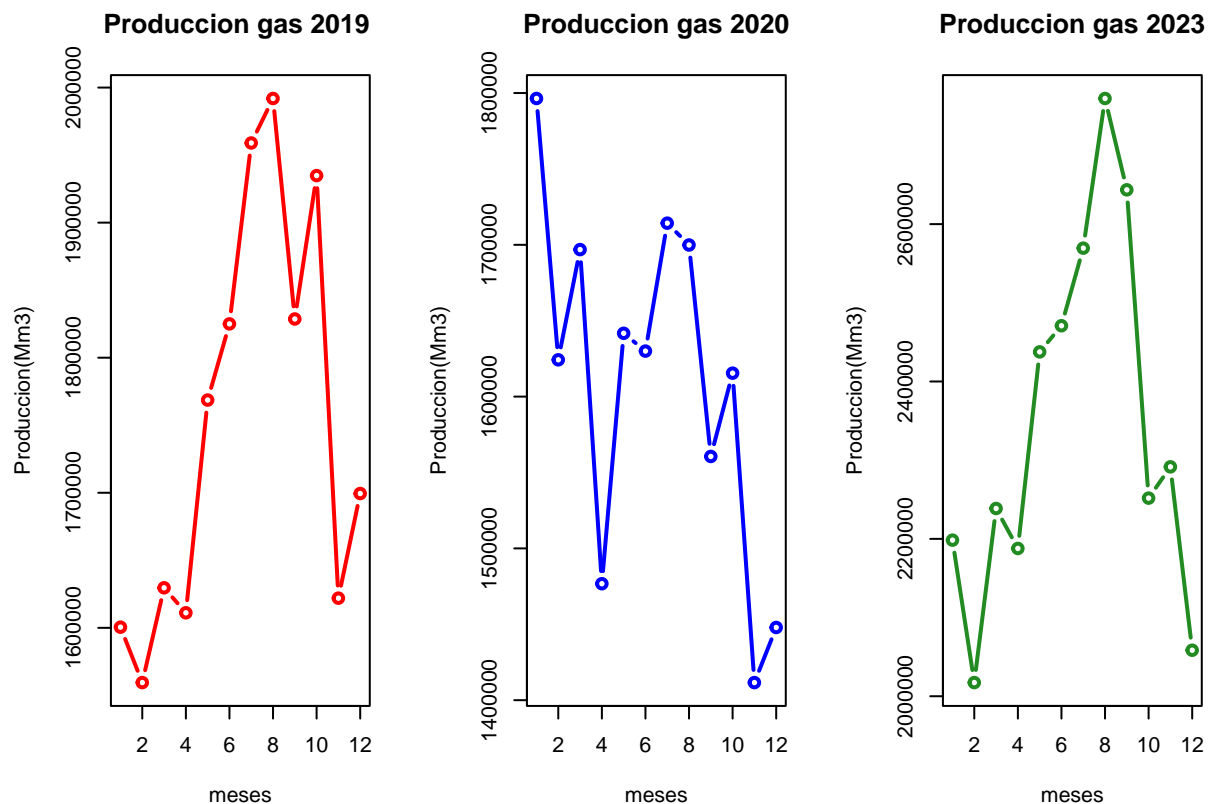
año2019 <- filtrardata%>%
  filter(anio == 2019)
meses<-group_by(año2019,mes)
produccion_mensualg<-summarise(meses,Promedio_mensual = mean(prod_gas),produccion_total= sum(prod_gas),)

año2020<- filtrardata%>%
  filter(anio == 2020)
meses<-group_by(año2020,mes)
produccion_mensual2020g<-summarise(meses,Promedio_mensual = mean(prod_gas),produccion_total= sum(prod_gas),)

año2023<- filtrardata%>%
  filter(anio == 2023)
meses<-group_by(año2023,mes)
produccion_mensual23g<-summarise(meses,Promedio_mensual = mean(prod_gas),produccion_total= sum(prod_gas),)

par(mfrow=c(1,3))
plot (produccion_mensualg$mes,produccion_mensualg$produccion_total, main= "Produccion gas 2019 ",xlab="meses",ylab="Produccion(Mm3)",col="red",type="o")
plot (produccion_mensual2020g$mes,produccion_mensual2020g$produccion_total ,main="Produccion gas 2020",xlab="meses",ylab="Produccion(Mm3)",col="blue",type="o")
plot (produccion_mensual23g$mes,produccion_mensual23g$produccion_total,main="Produccion gas 2023",xlab="meses",ylab="Produccion(Mm3)",col="green",type="o")

```



En estos graficos a diferencia del petróleo, que experimentó una abrupta caída en su producción durante el pico de la pandemia, el gas evidenció una mayor estabilidad, registrando un descenso menos significativo.

Pozos no convencionales en el la actulidad:

Para poder marcar los pozos más actuales en el mapa vamos a filtrar los datos del año 2024 y dado que tiene los pozos repetido por mes vamos a aplicar el filtro una segunda vez eligiendo el mes de enero.

Utilizamos la librería Leaflet para generar un mapa interactivo. En un principio, usamos la función sin realizar la clusterización, lo que provocaba que el mapa fuera pesado y no se pudiera realizar el zoom en las zonas importantes debido a la gran cantidad de pozos que había marcados. Para solucionar esto, usamos la función `markerClusterOptions()` que muestra agrupaciones de pozos que al hacer clic en ellas, muestra los pozos individuales en la zona. De este modo, se obtiene un mapa más fácil de cargar.

Otra función que le aplicamos son los popups, que al hacer clic en un pozo en particular muestran sus características principales.

```
library(dplyr)

pzos2024 <- produccion_de_pozos_de_gas_y_petrleo_no_convencional %>%

  filter(anio == 2024 & mes == 1)
pzos2024[1:10,32:36]

## # A tibble: 10 x 5
##   provincia coordenada_x coordenada_y tipo_de_recurso proyecto
##   <chr>          <dbl>         <dbl> <chr>          <chr>
## 1 Neuquén        -68.6         -38.3 NO CONVENCIONAL Sin Proyecto
## 2 Rio Negro      -67.8         -39.0 NO CONVENCIONAL Sin Proyecto
## 3 Rio Negro      -67.9         -39.0 NO CONVENCIONAL Sin Proyecto
## 4 Neuquén        -68.3         -38.7 NO CONVENCIONAL Sin Proyecto
## 5 Neuquén        -68.9         -38.3 NO CONVENCIONAL Sin Proyecto
## 6 Neuquén        -68.9         -38.3 NO CONVENCIONAL Sin Proyecto
## 7 Neuquén        -69.4         -38.6 NO CONVENCIONAL GAS PLUS
## 8 Neuquén        -68.9         -38.4 NO CONVENCIONAL Sin Proyecto
## 9 Neuquén        -68.9         -38.4 NO CONVENCIONAL Sin Proyecto
## 10 Neuquén       -68.9         -38.3 NO CONVENCIONAL Sin Proyecto

#“{r} library(leaflet)
leaflet(pzos2024) %>%
addProviderTiles(“OpenStreetMap.Mapnik”) %>%
addMarkers( lng = ~coordenada_x, lat = ~coordenada_y, popup = ~paste0( “ID Pozo:”, id_pozo, “”, “Em-
presa:”, id_empresa, “”, “Estado:”, tipo_estado, “”, “Profundidad:”, profundidad, “”, “Formación:”, formacion,
“”, “Petróleo (m3):”, prod_pet, “”, “Gas (Mm3):”, prod_gas ), clusterOptions = markerClusterOptions() )
%>%
setView(lng = -65, lat = -30, zoom = 4)

#““
```

Conclusión

Argentina se encuentra entre los países que han experimentado un mayor crecimiento en la producción de yacimientos no convencionales. En la actualidad, Neuquén es una de las provincias con la mayor cantidad de pozos perforados. Este auge se debe principalmente a la adopción de nuevas tecnologías y métodos que permiten perforar de manera más eficiente. Argentina se encuentra en una etapa crucial en el desarrollo de sus yacimientos no convencionales. Si se maneja de manera responsable y sostenible, Vaca Muerta tiene el potencial de transformar el panorama energético del país y contribuir a su crecimiento económico.