

Tarea 1 de Modelos Probabilistas Aplicados

Sector Construcción

5271

8 de septiembre de 2020

1. Origen de los datos

Los datos utilizados en este trabajo fueron obtenidos en el sitio <https://www.inegi.org.mx> que pertenece al Instituto Nacional de Estadística y Geografía (INEGI). Se escogió el apartado Construcción donde se muestra información sobre los principales resultados de las Empresas Constructoras, comprende unidades económicas dedicadas principalmente a la edificación; a la construcción de obras de ingeniería civil y a la realización de trabajos especializados de construcción. De este apartado se descargó el tabulado (Valor de producción generado por las empresas constructoras según el tipo de obra) en formato *csv*. En el cuadro 1 de la página 1 se muestra un fragmento de los datos descargados.

Cuadro 1: Fragmento del tabulado utilizado

Periodo	Año	Total	Edific	Agua_R_S	Elect_C	Transp	Petr_p	Otras_C
Enero	2006	23649637	11562266	1527733	1062658	5237209	2698498	1561273
Febrero	2006	23186956	11513255	1518386	808459	4901882	2931457	1513517
Enero	2007	27801512	13848554	1314745	1011836	6078270	3554327	1993780
Febrero	2007	27019544	13948606	949590	955818	6012856	3197691	1954983
Enero	2020	36031814	17703639	1363041	2418338	7684282	2265532	4596982
Febrero	2020	36626861	18064797	1186512	2189154	8031054	2539702	4615642

2. Análisis y tratamiento de los datos

Para el análisis de los datos del Valor de producción generado por las empresas constructoras según el tipo de obra, se tiene en cuenta que poseen una frecuencia de ocurrencia mensual, en el periodo comprendido entre los años 2006 y 2020. Por lo que se tratan los mismos como una serie de tiempo. El análisis será realizado en el programa R versión 4.0.2 [1] en el entorno de desarrollo Rstudio [2]. Para llevar al formato de serie de tiempo se realizó el código 2 de la página 1.

```
1 datos= read.csv("datos_new.csv", header = TRUE)
2 datos=datos[, -(1:2)]
3
4 Valor=ts(datos, start = c(2006,1), frequency = 12)
```

Series.R

2.1. Análisis

En el cuadro 2 de la página 2 se muestra un resumen de los resultados de varias funciones de ajuste de modelos aplicadas a los datos, Aunque la misma es bastante informativa, pero no tan amigable como una representación gráfica. Por lo anterior es conveniente utilizar el diagrama de caja y bigotes para una mejor comprensión de la información. En la figura 1 de la página 2 se muestra la relación entre el valor en miles de pesos corrientes y el tipo de obra.

Cuadro 2: resumen de varias funciones de ajuste de modelos aplicadas

	Edific	Agua_R_S	Elect_C	Transp	Petr_p	Otras_C
Min.	11513255	868901	769529	4901882	1717099	1277653
1st Qu.	14262184	1400904	1545408	7684372	2594559	2018676
Mediana	15600868	1596256	2267046	8712373	3007558	3458974
Media	16013749	1606719	2280334	8453855	3122855	3351165
3rd Qu.	17527312	1757474	2954560	9544727	3557916	4343621
Max.	22055350	2517616	5320331	1.1E+07	6061249	7027813

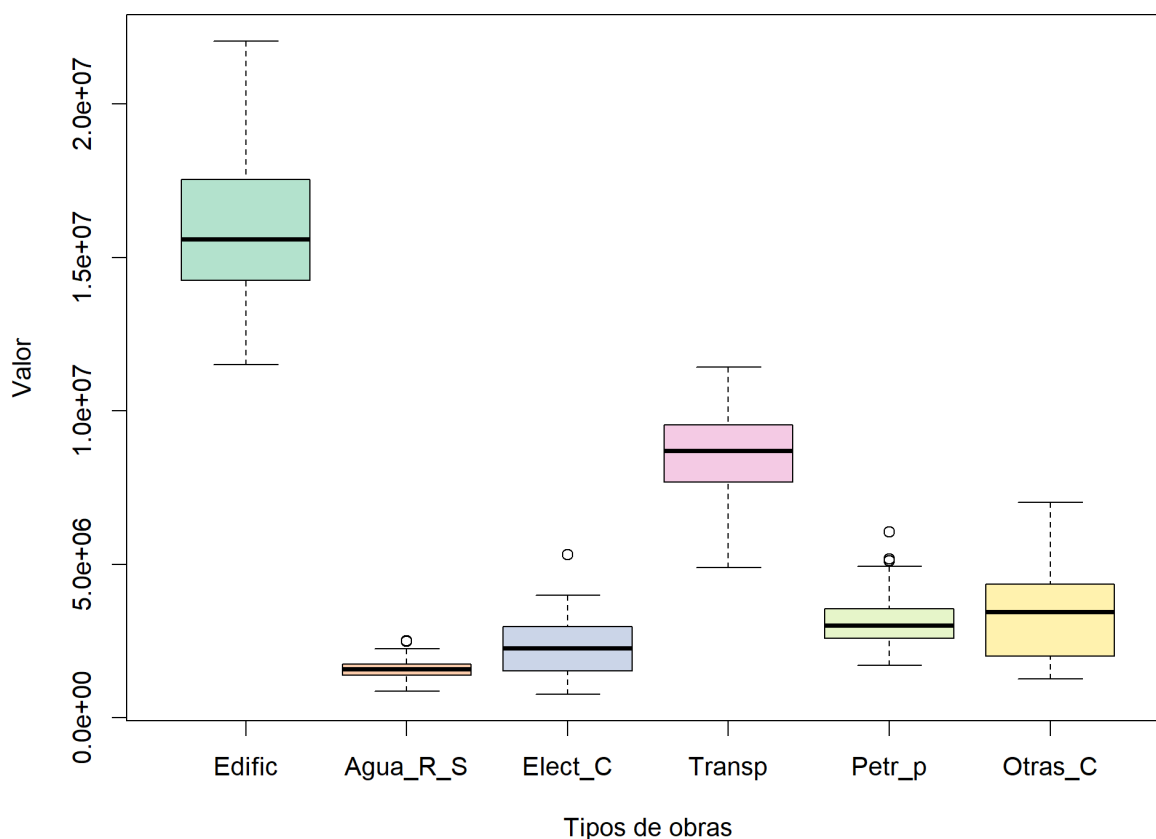


Figura 1: Diagrama de caja y bigotes que relaciona el valor en miles de pesos corrientes con los tipos de obras.

En el diagrama de caja y bigotes de la figura 1, se puede observar la diferencia entre los valores de los diferentes tipos de obras, además muestra que el tipo de obra Edificaciones es la de mayor valor de producción. Así como las de menor valor son la de Agua, riego y saneamiento y Electricidad y comunicaciones. Lo mismo se puede ver en la figura 2 de la página 3, donde se muestra una gráfica de secuencia que nos describe el comportamiento del valor de los diferentes tipos de obras a lo largo del tiempo. En la misma refleja que las obras en el sector Petróleo y petroquímica a partir de los años 2016 y 2017 tienen un menor valor que Electricidad y comunicaciones.

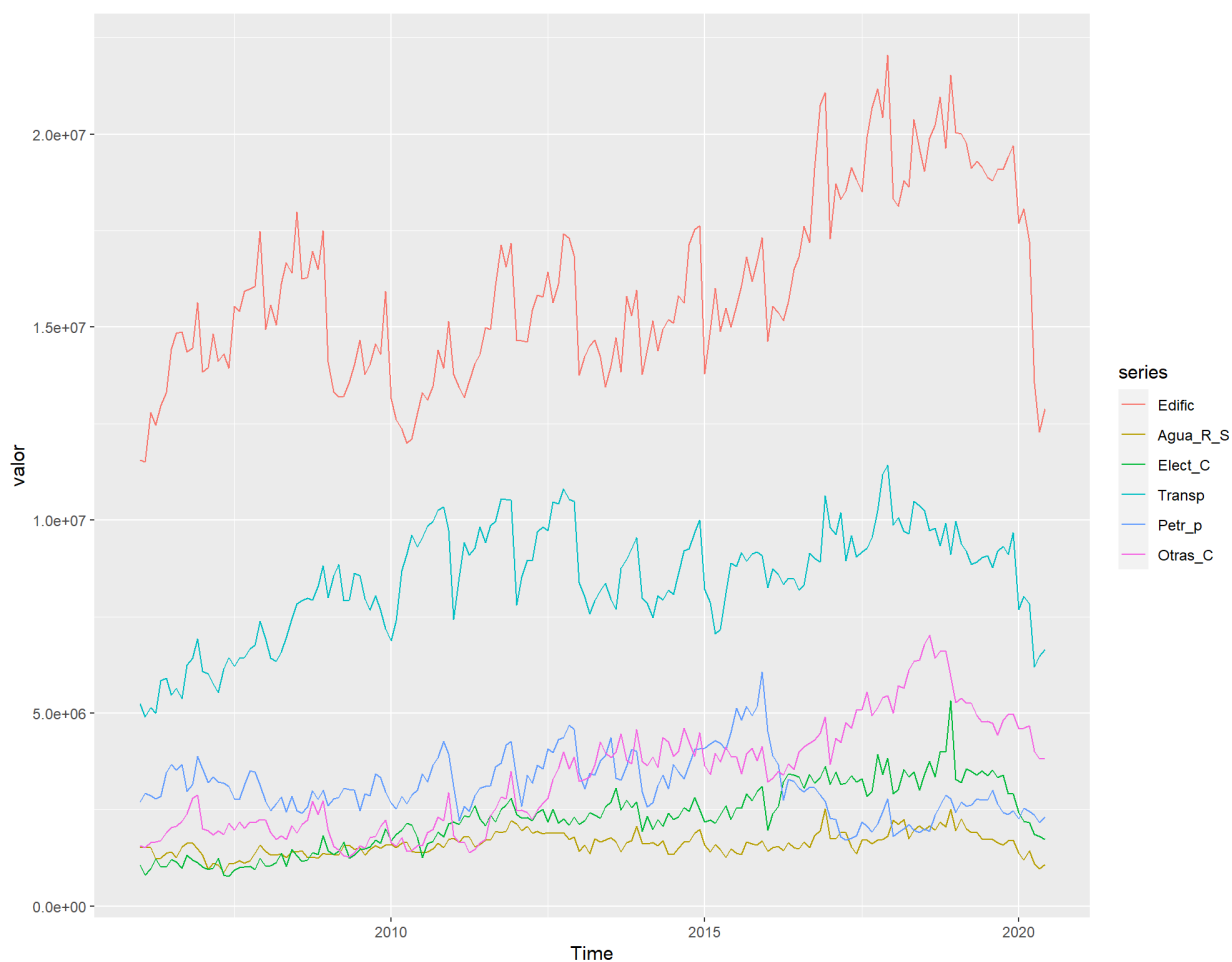


Figura 2: Diagrama de secuencia que muestra el comportamiento del valor de producción por tipo de obras a lo largo del tiempo.

Las figuras 3, 4, 5 6, 7, 8 de las paginas 4, 5, 6, 7, 8, 9 respectivamente, muestran los valores de producción de cada uno de los tipos de obra por meses. De la interpretación de estos diagramas se pueden inferir que los meses donde mayor valor se reporto fueron octubre, noviembre y diciembre de cada año.

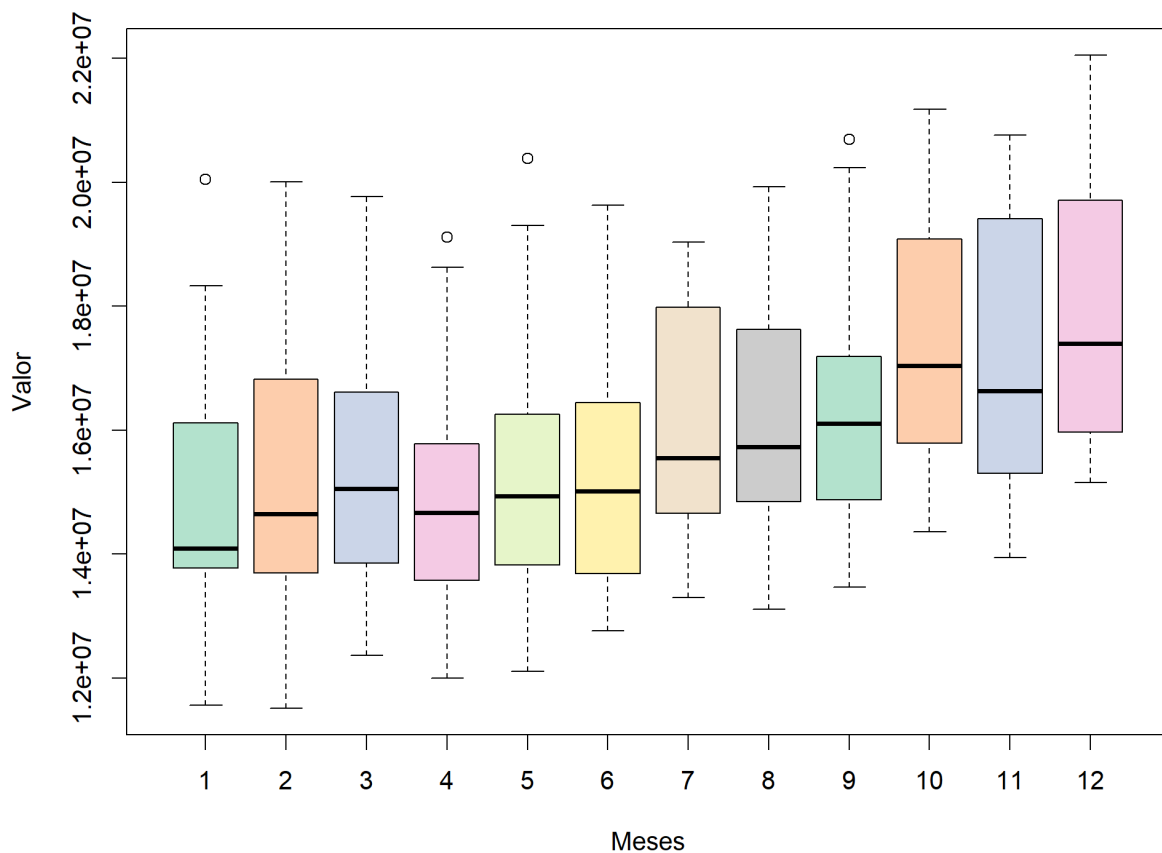


Figura 3: Diagrama de caja y bigotes del valor de producción del tipo de obra Edificación por meses.

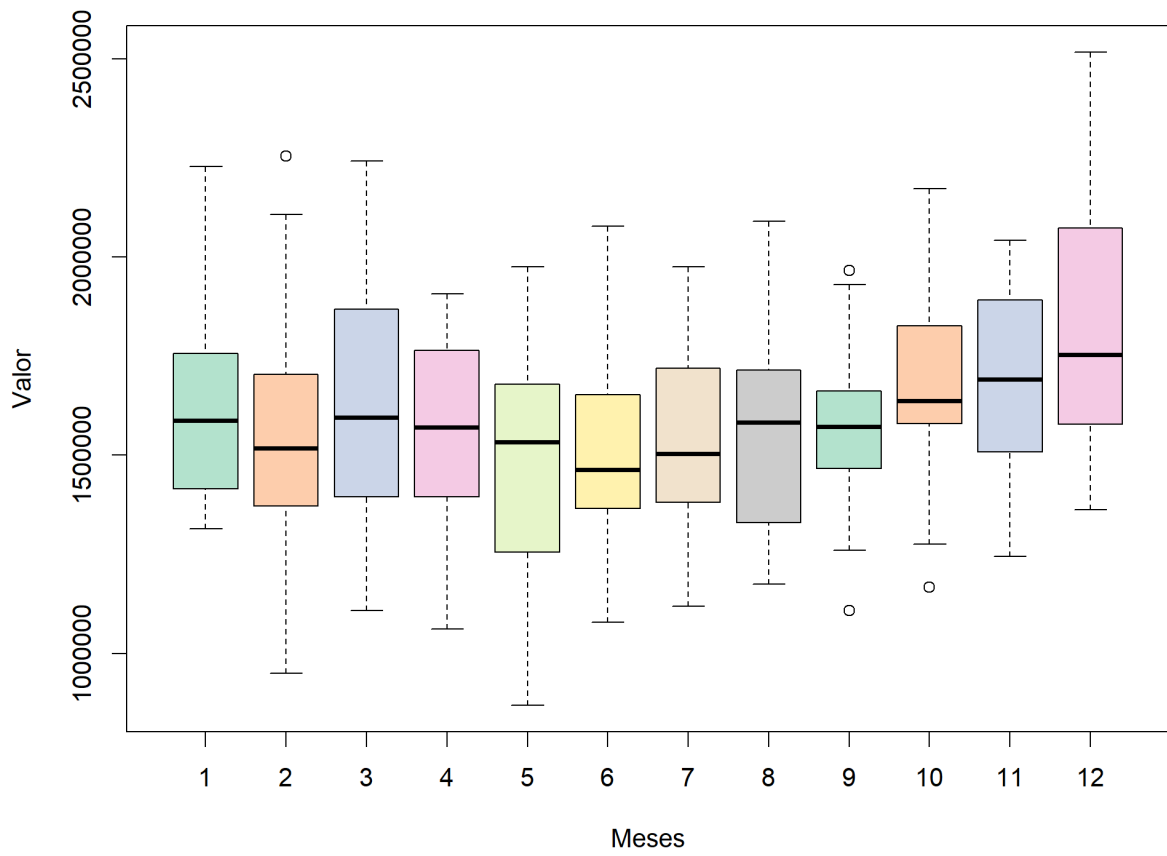


Figura 4: Diagrama de caja y bigotes del valor de producción del tipo de obra Agua, riego y saneamiento por meses.

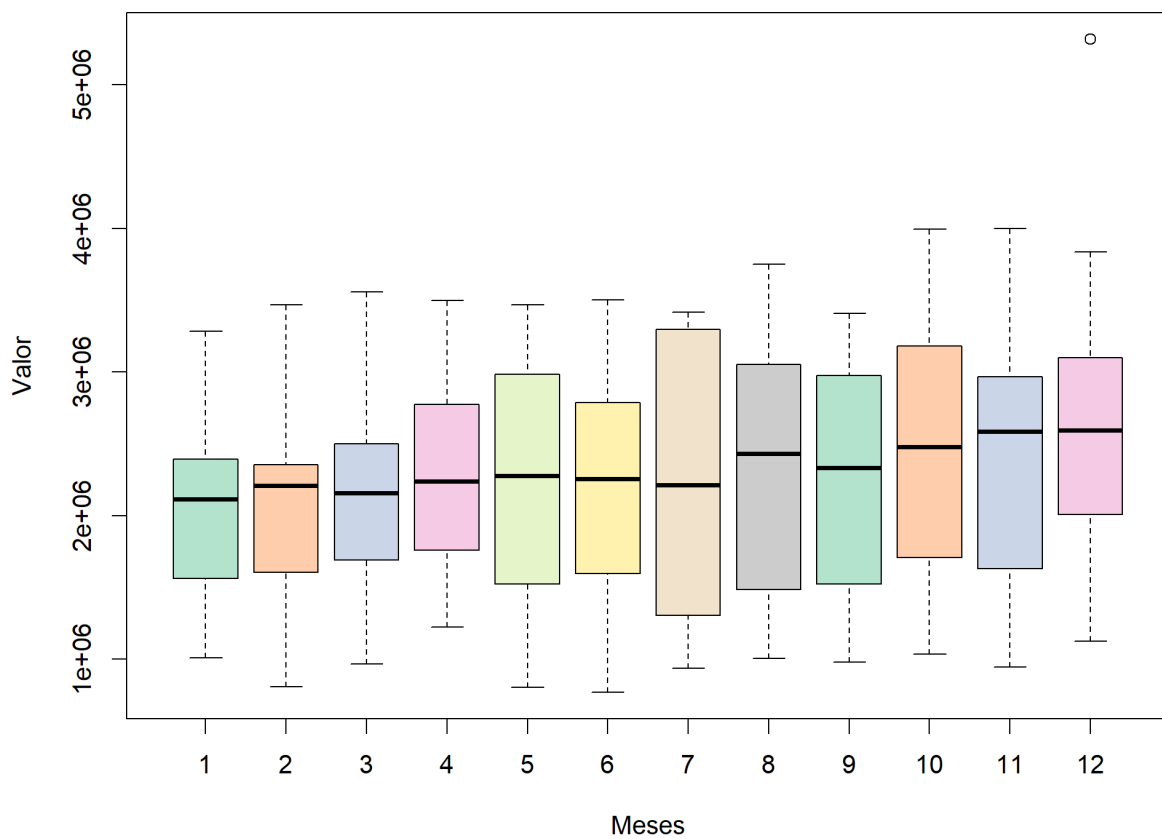


Figura 5: Diagrama de caja y bigotes del valor de producción del tipo de obra Electricidad y comunicaciones por meses.

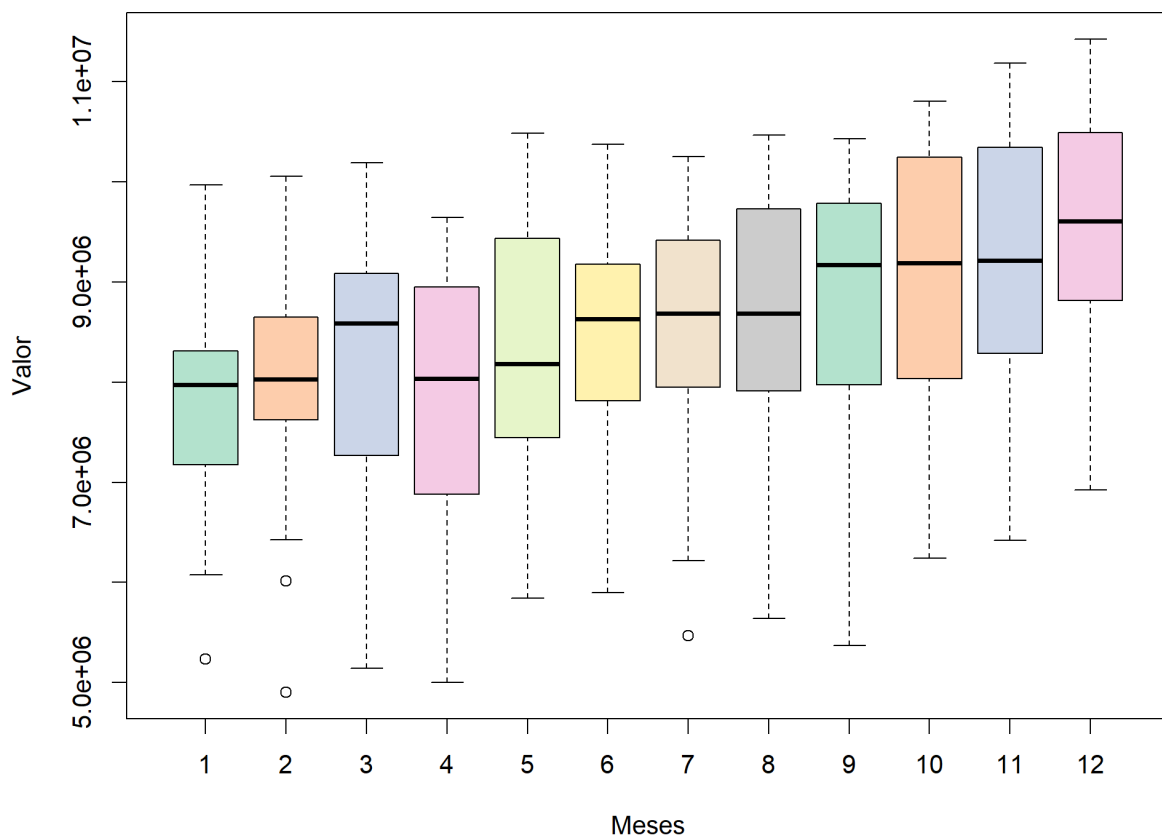


Figura 6: Diagrama de caja y bigotes del valor de producción del tipo de obra Transporte por meses.

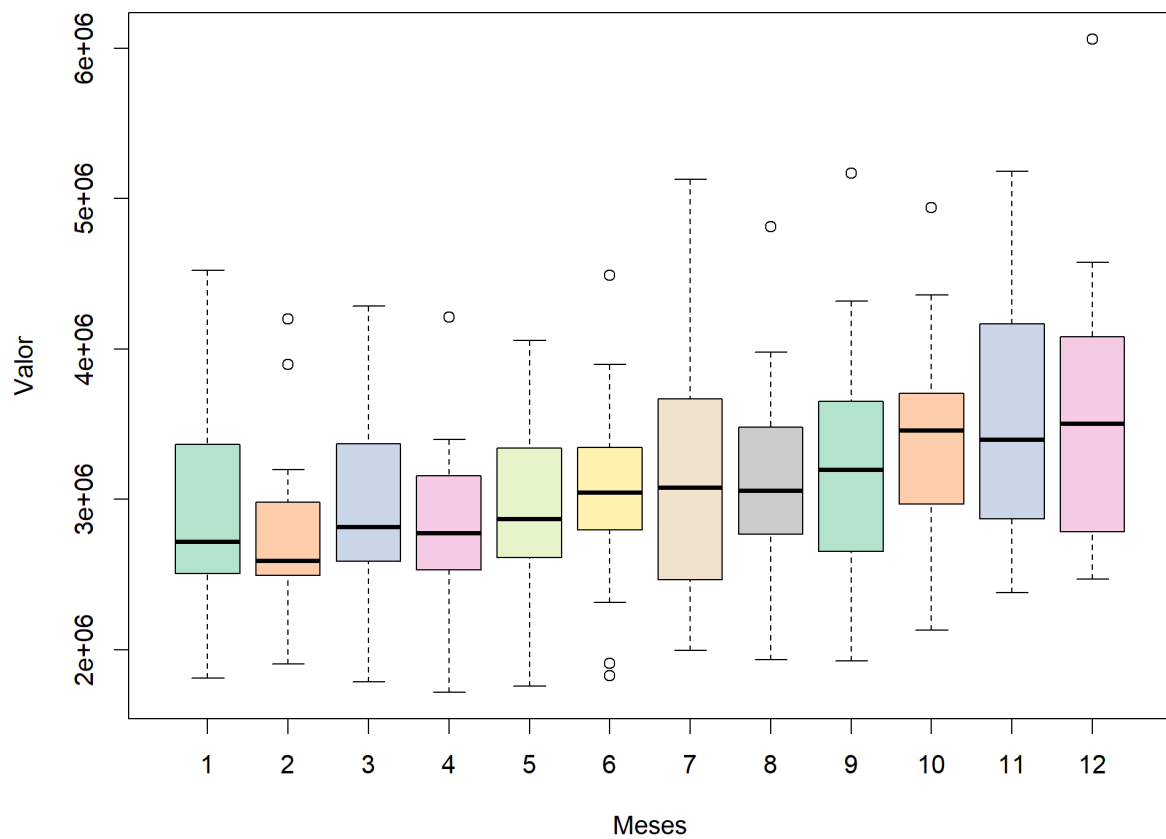


Figura 7: Diagrama de caja y bigotes del valor de producción del tipo de obra Petróleo y petroquímica por meses.

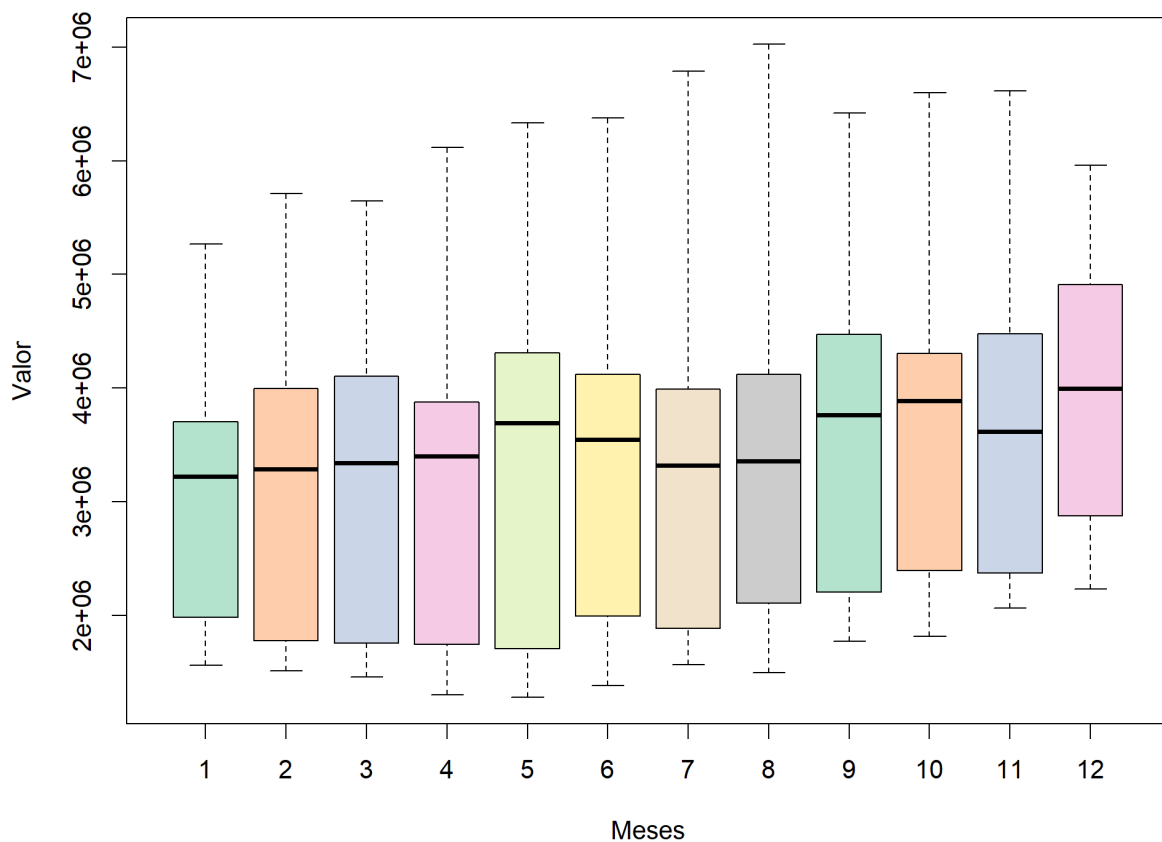


Figura 8: Diagrama de caja y bigotes del valor de producción del tipo de obra Otras obras por meses.

Para la realización del resumen y las gráficas se elaboró un el código de R 2.1 de la página 10, el código general se encuentra disponible en el repositorio https://github.com/Albertomnoa/Tareas_MPA/Tarea1

```

1 Sumario = summary(Valor)
2 capture.output(Sumario, file="Sumario.csv")
3
4 png(filename = "series.png",width = 2000, height = 1600, res =200)
5 autoplot(Valor, ts.colour = "blue", ts.linetype = "dashed")
6 dev.off()
7
8 png(filename = "boxplot.png",width = 2000, height = 1600, res =250)
9 boxplot(valor, col=palette("Pastel 2"),xlab="Tipos de obras",ylab = "Valor")
10 dev.off()
11
12 for (x in c(1:6)) {
13   png(filename = paste("boxplot",x,".png", sep=""),width = 2000, height = 1600, res
14     =250)
15   boxplot(Valor[,x]~cycle(Valor[,x]),col=palette("Pastel 2"),ylab="Valor",xlab = "
16     Meses")
17   dev.off()
18 }

```

Graficar.R

3. Conclusiones

- El tipo de obra con mayor valor de producción es la de Edificaciones y las de menor valor son Agua, riego y saneamiento y Electricidad y comunicaciones.
- A partir de los años 2016 y 20017 el tipo de obra Petróleo y petroquímicas bajo su valor de producción con respecto al sector Electricidad y comunicaciones.
- El valor de todos los tipos de obras presentan una caída a partir del año 2019 con respecto a la tendencia creciente de las mismas hasta el 20018.
- Los meses donde mayores valores de producción se reportaron fue en los del ultimo trimestre de cada año.

Referencias

- [1] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.
- [2] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, PBC., Boston, MA, 2020.