

## Trabajo práctico Nro. 3 - Análisis descriptivo

**Datos de permisos y registros de edificación** (fuente <https://datos.rosario.gob.ar/dataset/permisos-y-registros-de-edificación>)

Información desagregada sobre los permisos y registros de edificación desde el año 2005 hasta el 30/09/2018.

El conjunto de datos cuenta con 42935 registros. Los datos son provistos por la “Dirección General de Obras Particulares - Secretaría de Gobierno” de la municipalidad de Rosario.

### Variable “Tipo de trámite”

Variable cualitativa

**ON:** Obras Nuevas - Obra a construir (Permiso)

**M:** Obras Nuevas + Registro - Obra a construir con un sector a regularizar (Permiso)

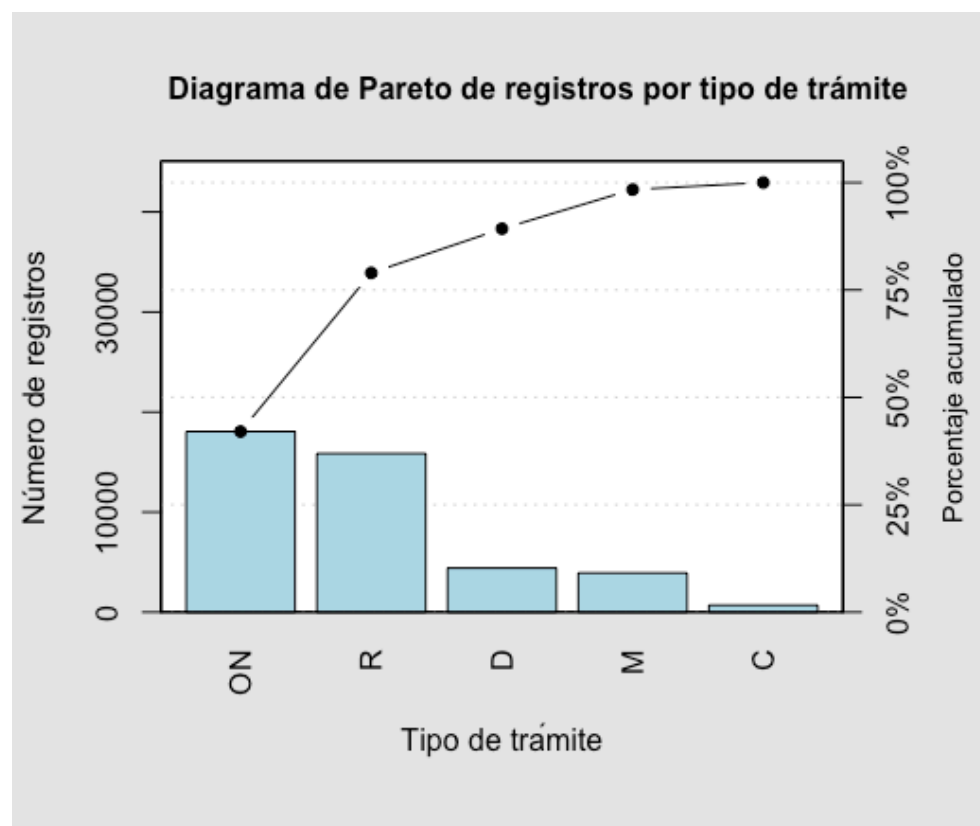
**R:** Registros - Registro de obra terminada

**C:** Circular 34 inc h - Trámite de Reposición de Planos al Archivo

**D:** Demoliciones

Diagrama de Pareto para Tipo de Trámite

En la figura 1 se puede observar que los tipos de trámite “Obras Nuevas” (ON) y “Obras Nuevas + Registro” (M) representan más del 75% de los registros.



**Figura 1** Diagrama de Pareto de registros por tipo de trámite. En el gráfico se puede observar que los tipos de trámite ON y R representan más del 75% de los registros.

**Tabla 1 Distribución de los registros según el tipo de trámite.**

En la tabla se confirma lo observado en la figura 1, los tipos de trámite ON y R tiene las mayores frecuencias.

Tipo de trámite	Frecuencia absoluta	Frecuencia relativa	Frecuencia relativa porcentual
ON	18052	0.42	42.05
M	3919	0.09	9.13
R	15858	0.37	36.93
D	4410	0.10	10.27
C	696	0.02	1.62
Total	42935	1.00	100

## Variable “Total metros cubiertos (en m<sup>2</sup>)”

Variable cuantitativa continua.

Medidas centrales

**Media aritmética:** 302.43 m<sup>2</sup>

**Mediana:** 102 m<sup>2</sup>

**Moda:** 59 m<sup>2</sup>

**Minimo:** 1 m<sup>2</sup>

**Maximo:** 118083 m<sup>2</sup>

**Rango:** 118082 m<sup>2</sup>

**Varianza:** 1747872

**Desvío Estándar:** 1322.07 m<sup>2</sup>

**Cuartiles:**

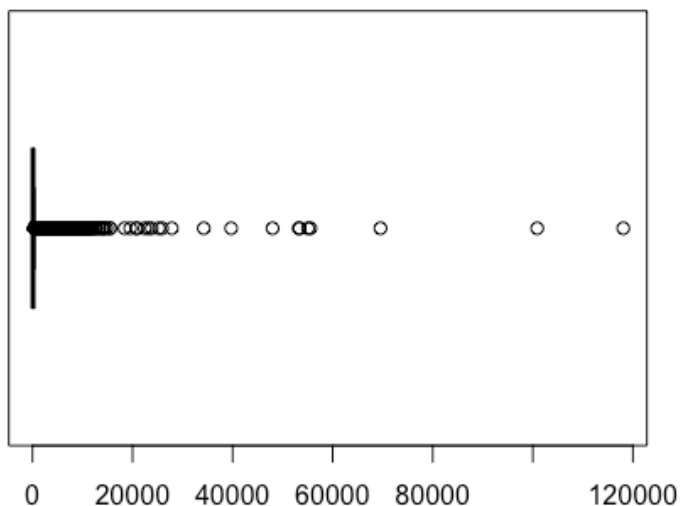
0%	25%	50%	75%	100%
1	52	102	208	118083

Observando las medidas centrales, se puede encontrar que la media no es representativa dada la presencia de valores extremos. Por lo tanto la mediana (102 m<sup>2</sup>) es una medida mas representativa de la distribución del total de metros cuadrados cubiertos, en la muestra.

### **Diagramas de caja**

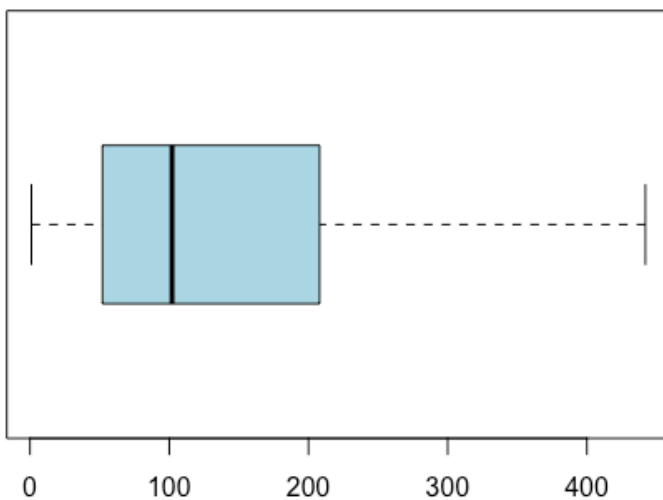
Se puede observar mayor variabilidad en los metros cubiertos, para los tipos de trámite de obras nuevas (ON), como se puede comprobar en la figura 5. También este tipo de trámite es el que posee mayor cantidad de valores extremos, junto con el tipo de trámite Circular 34 (C).

#### **Total metros cuadrados cubiertos**

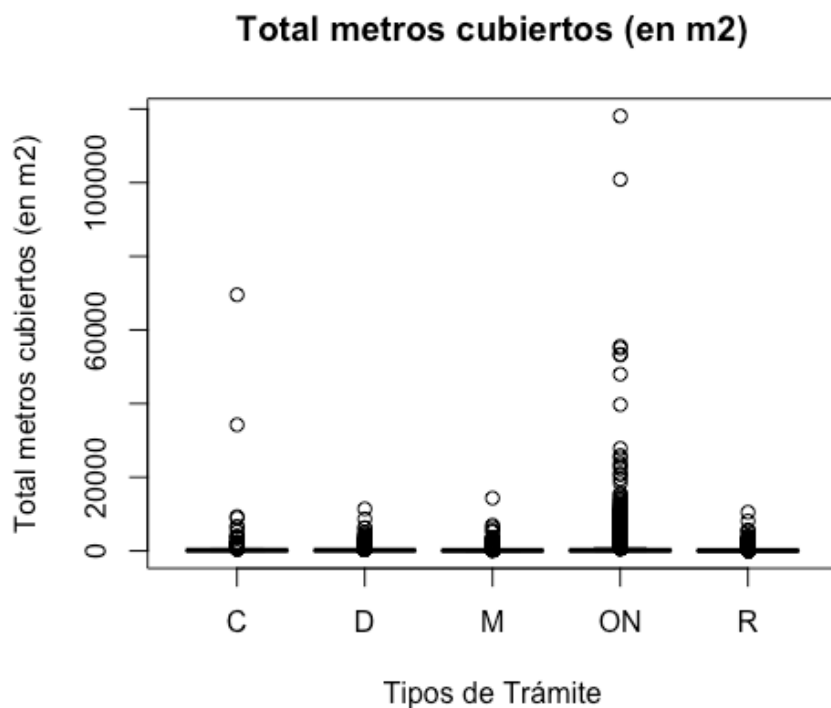


**Figura 2 Diagrama Box plot con la totalidad de registros.** Se pueden visualizar valores extremos muy grandes, lo que dificulta la visualización del boxplot de manera apropiada.

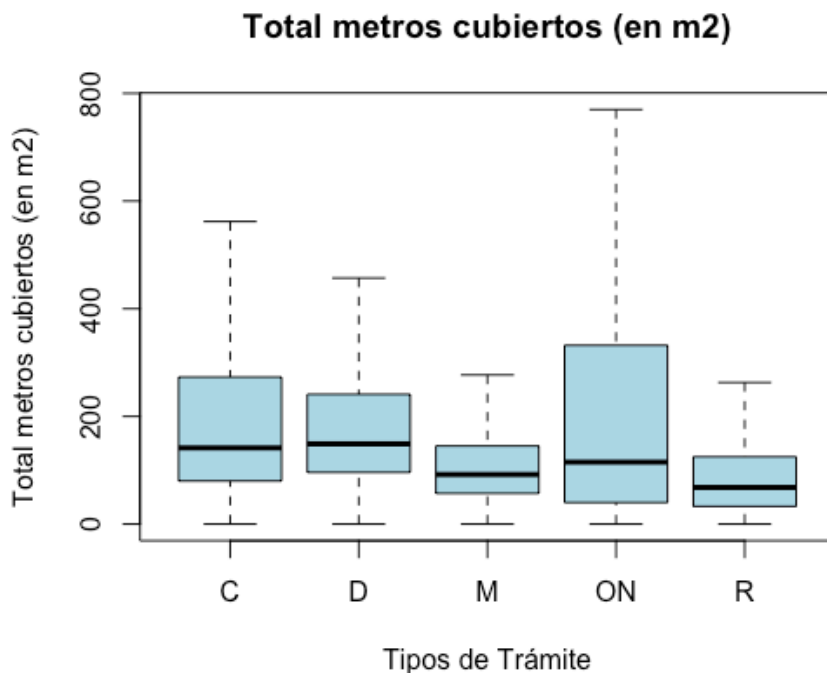
#### **Total metros cubiertos (en m2)**



**Figura 3 Diagrama Box plot excluyendo valores extremos.** En este diagrama no se incluyen los valores extremos especificando el parámetro OUTLINE de la función boxplot de R.



**Figura 4 Diagrama Box plot agrupado por tipo de trámite.** En este diagrama se puede visualizar que los trámites "C" y "ON" poseen los valores muy extremos. Esos valores podrían ser revisados y verificar si hubo un problema en la carga de datos.



**Figura 5 Diagrama Box plot agrupado por tipo de trámite.** En este diagrama se especifica el parámetro OUTLINE de R para excluir los valores extremos y poder visualizar mejor los boxplots. Se observa que el tipo de trámite "ON" tiene mayor variabilidad.