




Organización de Datos 75.06/95.58
Trabajo Práctico N° 1

GRUPO: 1

Apellido y Nombre	Padrón	Correo electrónico
Funes Federico	98372	fede.funes96@gmail.com
Caceres Julieta	96454	julieta.agustina.caceres@gmail.com
Lasanta Tito	99162	titonoolvida@gmail.com
Almejún Santiago	99441	santi.alme@gmail.com

Github: 

<https://github.com/fedefunes96/TpDatos>

Índice

1. Introduction	3
2. Cleanup previo	3
3. Avisos	3
3.1. Distribución de las postulaciones desde la publicación de un aviso	3
3.2. Top de empresas según cantidad de postulaciones y de avisos .	4
3.3. Características de los avisos (Descripción de los mismos) . . .	6
3.3.1. Top de características de un aviso según cantidad de avisos	7
3.3.2. Capacitación	8
3.3.3. Buen clima laboral	9
3.3.4. Disponibilidad horaria	10
3.3.5. Disponibilidad para viajar	11
3.3.6. Estudiantes	12
3.3.7. Experiencia	13
3.4. Postulaciones segun rubro del anuncio	14
3.4.1. Tipos de trabajo con mayor cantidad promedio de pos- tulantes	15
3.4.2. Tipos de trabajo con menor cantidad promedio de pos- tulantes	16
3.4.3. Conclusión	17
4. Postulantes	18
4.1. Cantidad de postulantes divididos por edad	18
4.2. Cantidad de postulantes divididos por nivel educativo	19
5. Evolución Temporal	20
5.1. Postulaciones según hora del día	20
5.2. Vistas según hora del día	21
5.3. Postulaciones según día de la semana	21
5.4. Vistas según día de la semana	22
5.5. Postulaciones según día del mes	23
5.5.1. Postulaciones según día del mes - Enero	23
5.5.2. Postulaciones según día del mes - Febrero	24
5.6. Vistas según día del mes	25
5.7. Conclusión sobre el aspecto temporal	26

6. Nivel laboral y Nivel educativo.	27
6.1. Postulaciones por nivel educativo.	27
6.2. Publicaciones por nivel laboral.	31
6.2.1. Postulantes a Senior/Semisénior.	31
6.2.2. Postulaciones para Junior.	37
6.2.3. Postulaciones para Jefe.	38
6.2.4. Postulaciones para Gerencia / Alta Gerencia / Dirección.	39
6.3. Conclusión sobre nivel laboral y educacional.	40

1. Introduction

El objetivo de este informe es mostrar el trábajo de análisis exploratorio realizado sobre el set de datos proporcionado por la empresa Navent de su sitio <https://www.zonajobs.com.ar/>. Los datos constan de los registros de avisos de búsquedas laborales para los primeros meses del año 2018.

2. Cleanup previo

El set de datos provisto consiste en varios archivos CSV, dos sobre los datos de los postulantes, dos sobre datos de los avisos y otros dos con relaciones entre ambos (vistas y postulaciones). En general, no hubo que hacer mucho filtrado sobre los datos, ya que los mismos venían, en la mayoría de los casos, válidos, lo único que se eliminó fueron datos sobre los avisos que casi no tenían valores. Para el resto del análisis, usamos dos datasets para los detalles de los avisos, uno que contiene información sobre las postulaciones (Se filtraron avisos sin postulaciones y aquellos que no duraran al menos 1 semana), y uno con los detalles en general para los avisos. La motivación de tener 2 datasets para los detalles de los avisos es que los avisos filtrados por postulaciones eliminan bastantes datos, y para los análisis que no requieren las postulaciones, es mejor tener la mayor cantidad de datos.

3. Avisos

En ésta sección, se harán análisis sobre los avisos, ya sean postulaciones en los mismos, distribuciones, características, etc...

3.1. Distribución de las postulaciones desde la publicación de un aviso

Para analizar la distribución de las postulaciones a lo largo del tiempo para un aviso desde que se publica, es necesario tener la fecha inicial y final de mismo. Como estos datos no fueron provistos, consideramos que la primera vez que algún usuario ve un aviso o se postula (difieren las fechas de los mismos, por eso consideramos también las postulaciones y nos quedamos con el mínimo entre ambas), es el día que se publica el aviso, y para la fecha final, la última vez que alguien vio el aviso o se postulo al mismo (quedándonos esta vez con el máximo entre ambas). De manera que ahora es posible ver como se distribuyen las postulaciones:

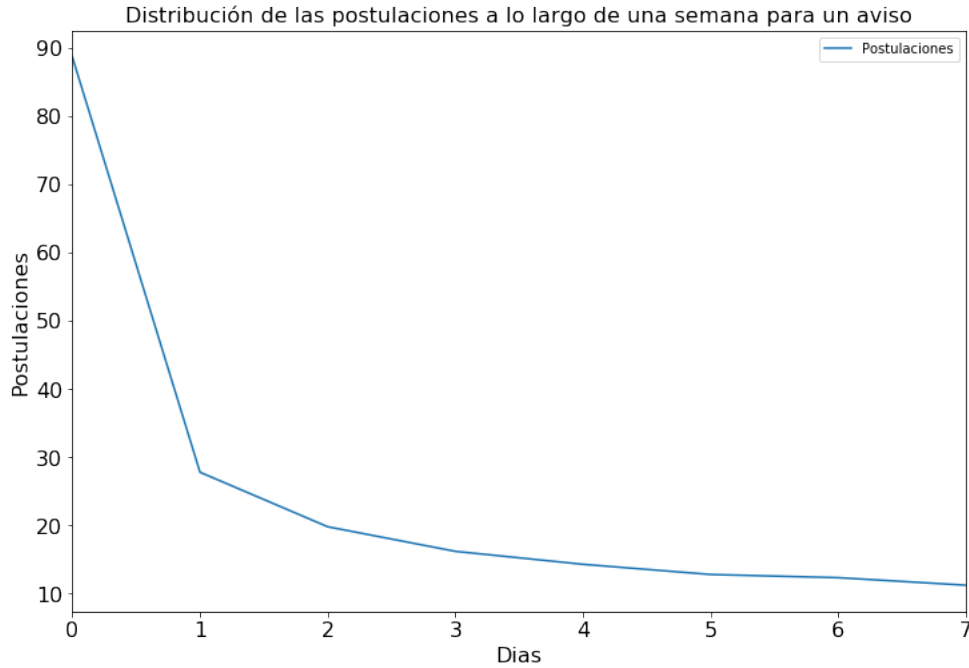


Figura 1: Distribución de la cantidad de postulaciones para un aviso desde su publicación

Se observa que los primeros días de un aviso son cruciales, dado que el flujo de postulantes es mayor en este momento, mientras que al pasar el tiempo la cantidad de nuevas postulaciones va tendiendo a 0. Este fenómeno es característico de redes sociales, por ejemplo en twitter, un tweet recibe la mayoría de sus comentarios o vistos en los primeros días.

3.2. Top de empresas según cantidad de postulaciones y de avisos

Para considerar el top según la cantidad de postulaciones y avisos, nos quedamos con aquellos avisos que duraron al menos una semana desde que se publicaron, ya que como se ha visto en la Figura 1, la mayoría de las postulaciones para un aviso se producen en los primeros días desde su publicación. Como solo nos dan datos sobre las postulaciones y avisos entre el 15-01-2018 y el 01-03-2018, analizamos en base a esas fechas, también, consideramos empresas que tengan al menos 5 avisos.

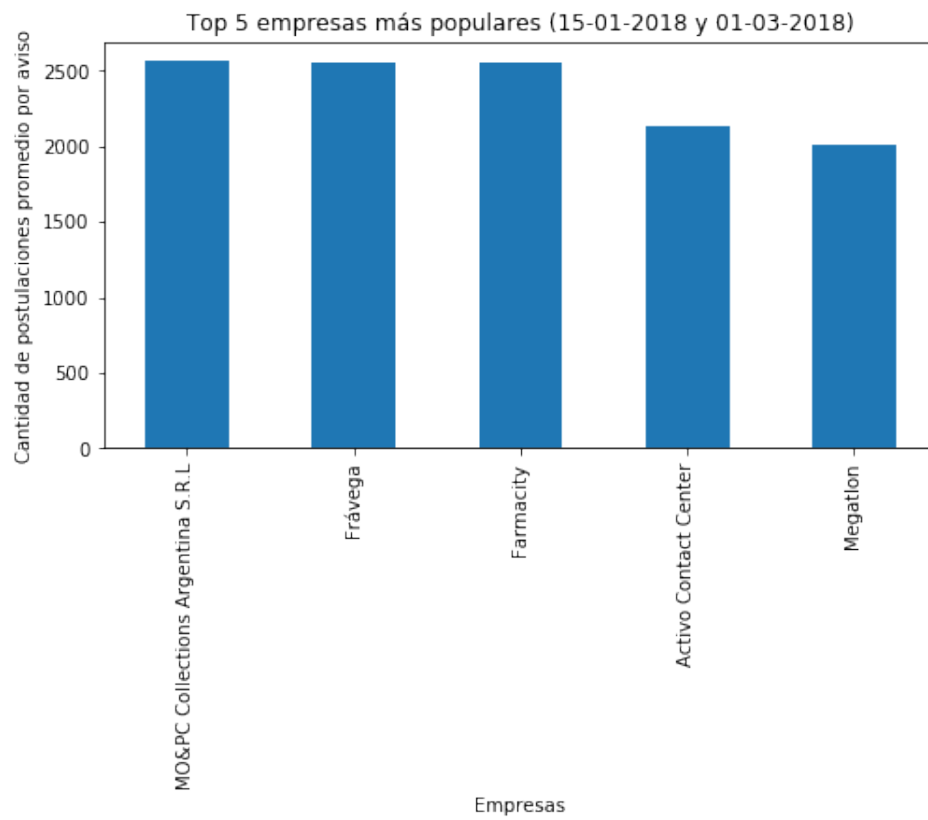


Figura 2: Top 5 empresas según postulaciones promedio a sus avisos

Se observa que las empresas más populares son, empresas grandes, conocidas y nacionales.

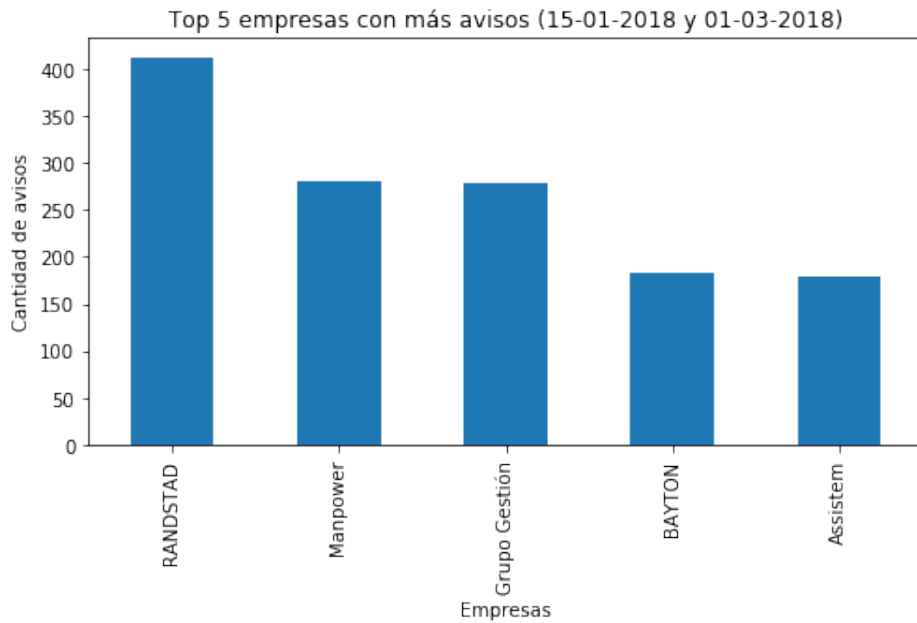


Figura 3: Top 5 empresas según cantidad de avisos

3.3. Características de los avisos (Descripción de los mismos)

Como los detalles de los avisos proveen una columna de descripción del mismo, parseamos del mismo palabras claves que podrían determinar cambios en la cantidad de postulaciones promedio a un aviso, cambios en la edad promedio de las personas que se postulan a un aviso, sus estudios, etc. Los términos claves determinados fueron los siguientes:

- **Web:** Si un aviso posee página web.
- **Crecimiento:** Si un aviso ofrece crecimiento laboral y/o progreso.
- **Capacitación:** Se ofrece capacitación.
- **Clima laboral:** Se ofrece un buen clima laboral.
- **Multinacional:** Si el aviso forma parte de una empresa multinacional.
- **Incorporación inmediata:** Se ofrece incorporación inmediata al trabajo.
- **Disponibilidad horaria:** Se exige disponibilidad horaria para trabajar fines de semana, sábados, feriados y horarios rotativos.

- **Secundario:** Si el aviso exige secundario completo.
- **Terciario:** Si el aviso exige terciario completo.
- **Inglés:** Si el aviso exige conocimientos básicos de inglés.
- **Estudiantes:** Si el aviso pide estudiantes universitarios.
- **Conocimientos informáticos:** Si el aviso pide conocimientos básicos del paquete office (Word, Excel, etc).
- **Experiencia:** Si el aviso pide experiencia previa en el rubro.
- **Proactividad:** Si el aviso pide postulantes que sean proactivos.

3.3.1. Top de características de un aviso según cantidad de avisos

Se procede a analizar cuales son las características de los avisos que son más frecuentes:

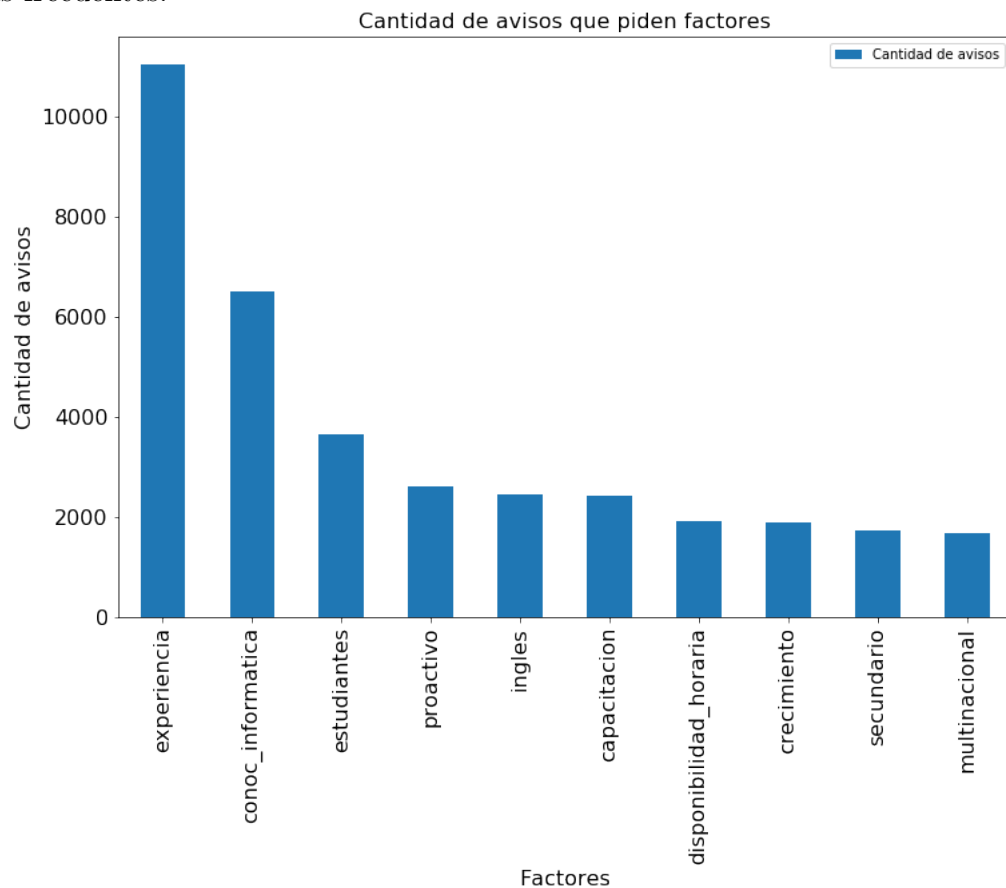


Figura 4: Top 10 características de avisos

3.3.2. Capacitación

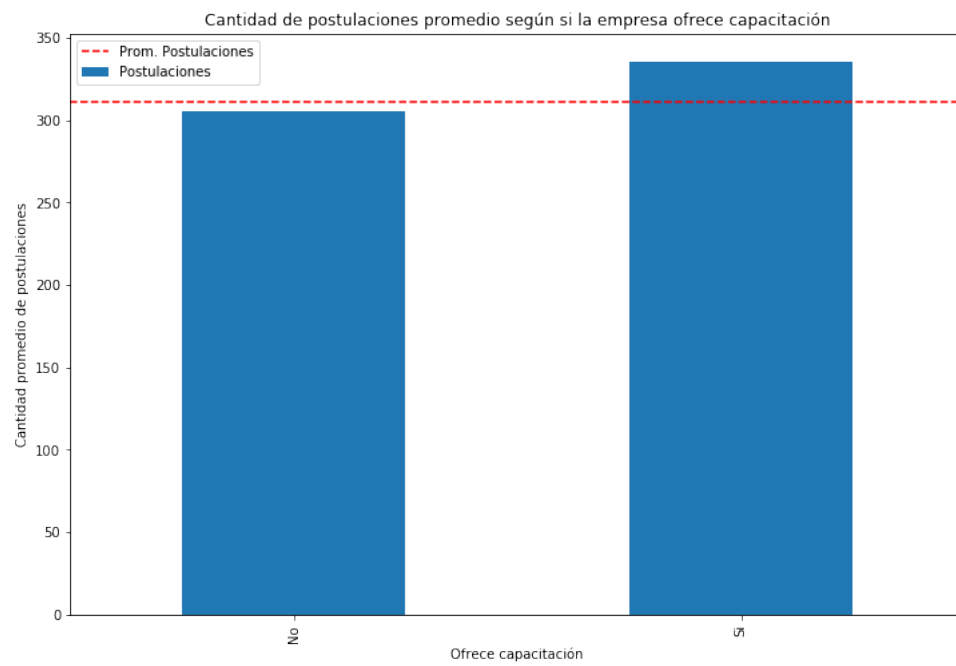


Figura 5: Promedio de postulaciones según si el aviso ofrece capacitación

3.3.3. Buen clima laboral

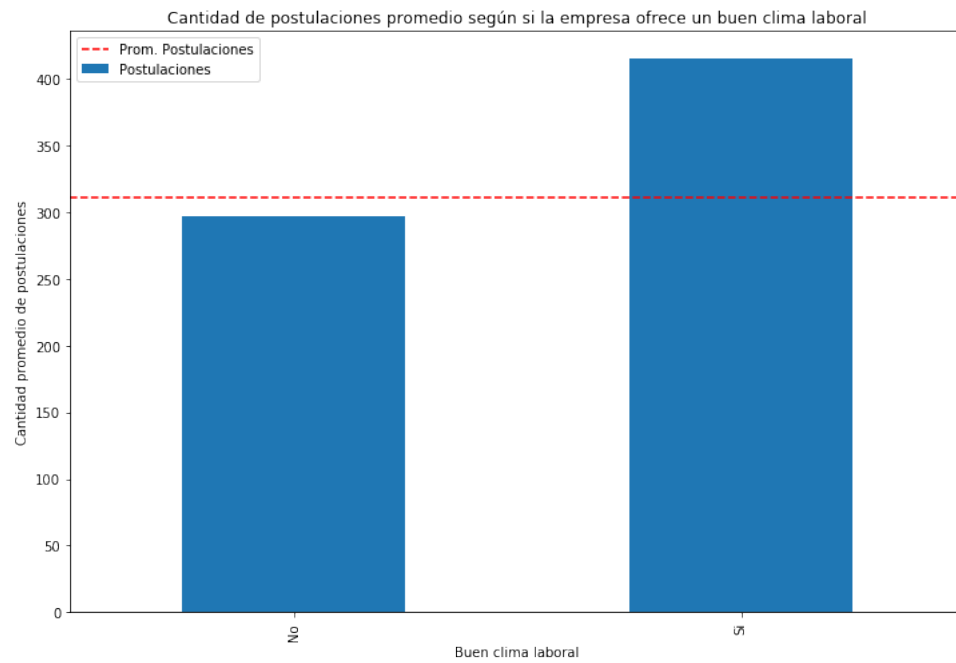


Figura 6: Promedio de postulaciones según si el aviso ofrece buen clima laboral

3.3.4. Disponibilidad horaria

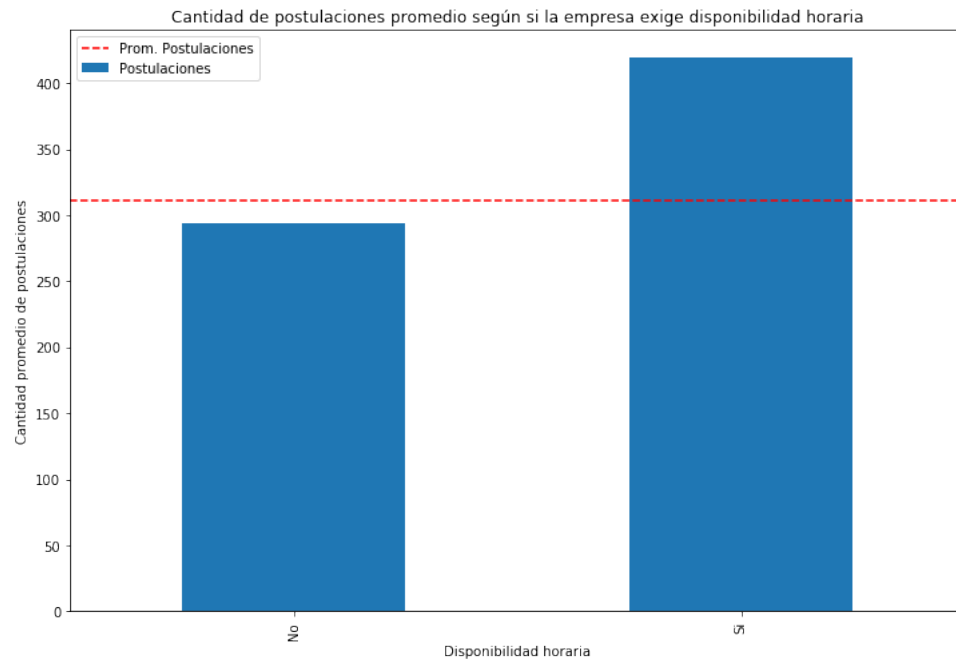


Figura 7: Promedio de postulaciones según si el aviso exige disponibilidad horaria

3.3.5. Disponibilidad para viajar

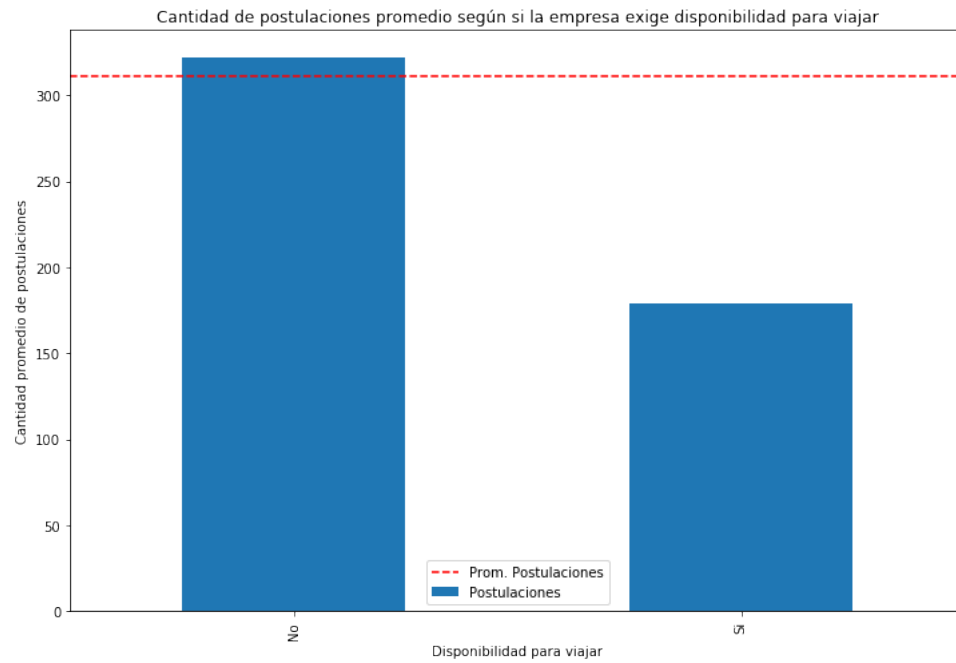


Figura 8: Promedio de postulaciones según si el aviso exige disponibilidad para viajar

3.3.6. Estudiantes

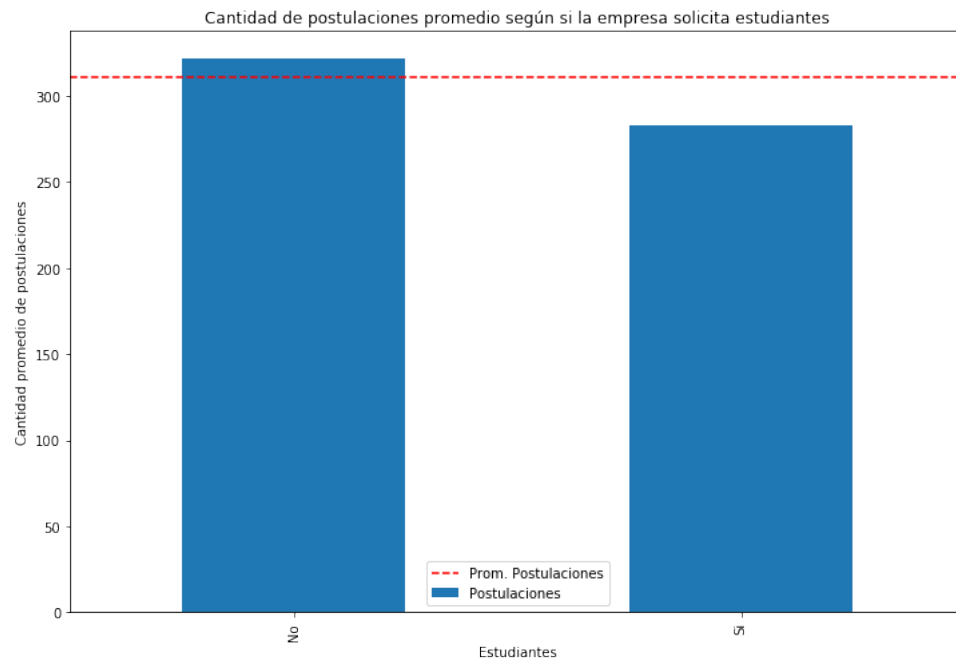


Figura 9: Promedio de postulaciones según si el aviso exige estudiantes universitarios

3.3.7. Experiencia

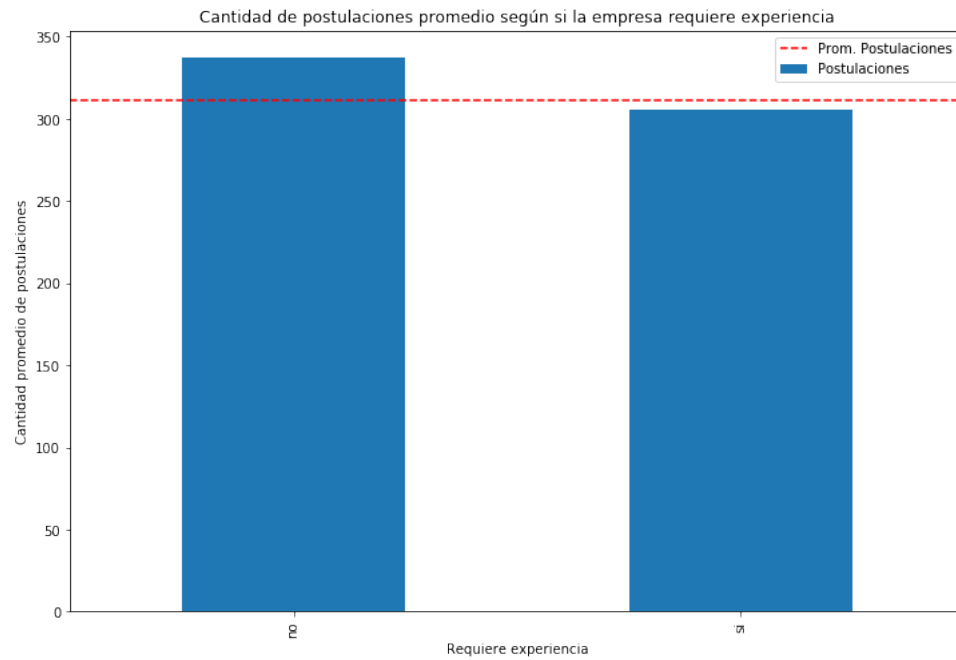


Figura 10: Promedio de postulaciones según si el aviso exige experiencia previa

3.4. Postulaciones segun rubro del anuncio

¿Depende la cantidad de postulantes que un aviso recibe del tipo de trabajo del que se trata? Es esperable que si, ¿pero cuanto?

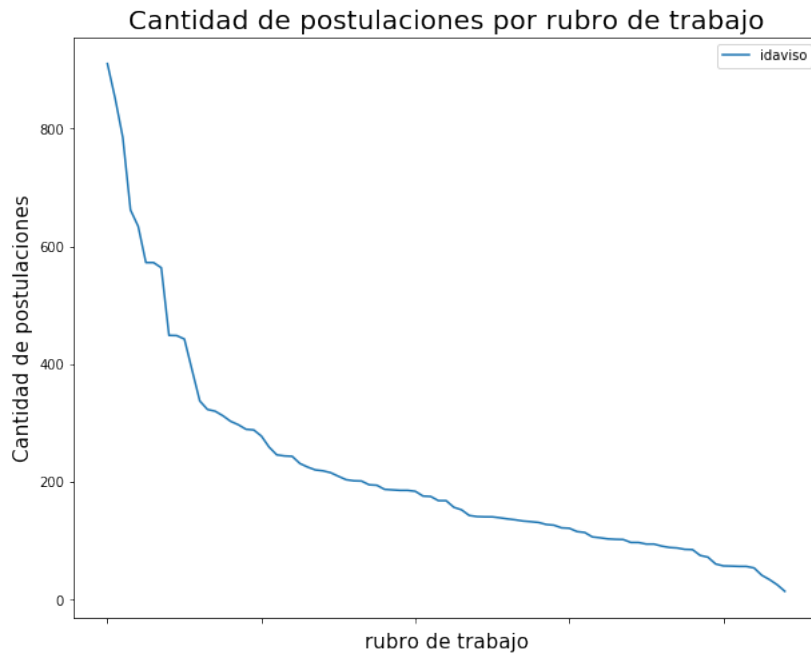


Figura 11: Promedio de postulaciones según el tipo de trabajo

Se ve que, dependiendo solamente de su rubro, un aviso puede llegar a obtener 10 veces más postulantes. ¿Se pueden relacionar los tipos de trabajo según la cantidad de postulantes que reciben? Hacemos zoom a las 2 puntas del grafico:

3.4.1. Tipos de trabajo con mayor cantidad promedio de postulantes

Trabajos con mayor cantidad de postulaciones por rubro de trabajo

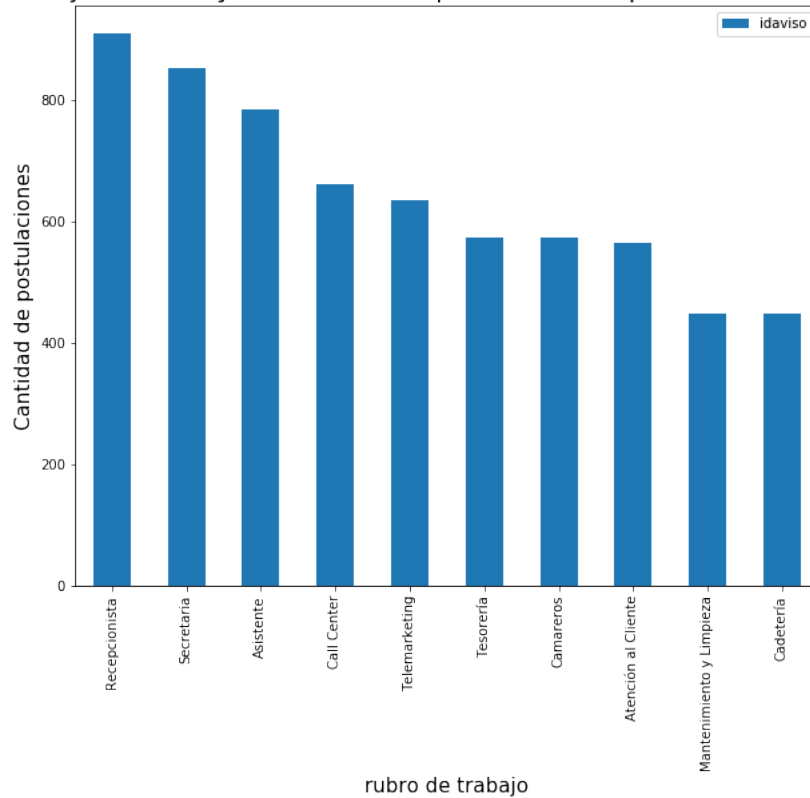


Figura 12: Tipos de trabajo con mayor promedio de postulaciones

Se puede ver que los tipos de trabajos con mayor cantidad de postulantes promedio son aquellos que no requieren conocimientos previos, particularmente la mitad se tratan de trabajos de escritorio y con excepciones se trata de trabajos del ámbito administrativo o de atención al cliente.

3.4.2. Tipos de trabajo con menor cantidad promedio de postulantes

Trabajos con menor cantidad de postulaciones por rubro de trabajo

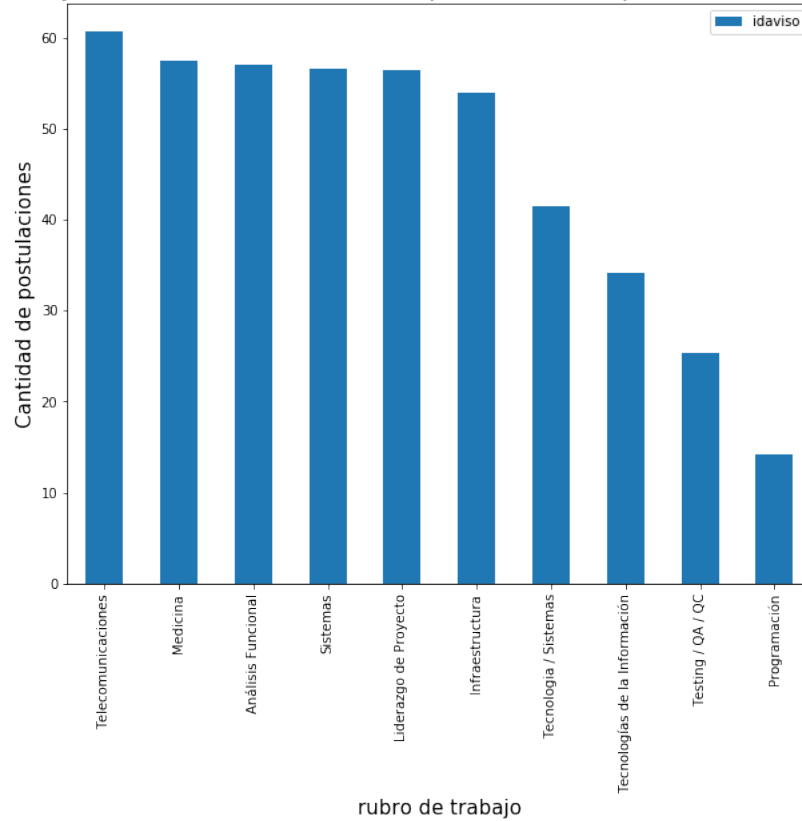


Figura 13: Tipos de trabajo con menor cantidad promedio de postulaciones

En este ejemplo, que presenta las categorías de trabajo con menor cantidad promedio de postulaciones, se ve claramente una agrupación de los tipos de trabajo. En contraste con el resultado previos todos las categorías presentadas requieren conocimiento u estudios previo, sin embargo, lo mas notable es que 8 de los 10 nombres son del área de Tecnología, sistemas o desarrollo de software.

Trabajos con menor cantidad de postulaciones por rubro de trabajo

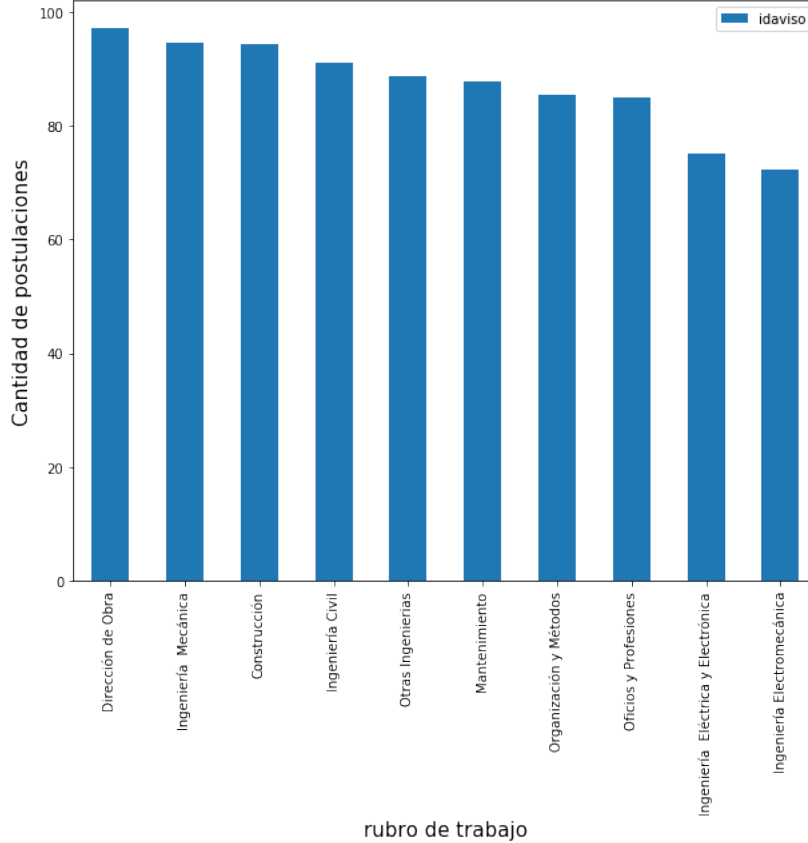


Figura 14: Tipos de trabajo con menor cantidad promedio de postulaciones (posiciones(10-20))

Si se observan las posiciones anteriores a las del gráfico previo, se encuentran nuevamente temas en común, se puede destacar que ingeniería aparece en la mitad, o que un tercio se trata de construcciones.

3.4.3. Conclusión

A partir de estos ejemplos se puede argumentar que es muy probable que los postulantes se postulen a avisos que entre sí tengan cantidades similares de postulantes. Una posible explicación es que, considerando que la oferta de tipos de postulantes no es uniforme, si no que al agruparlos por una característica, por ejemplo su nivel educativo, hay diferentes cantidades por grupo. Entonces, si el interés de los postulantes hacia un anuncio depende de a qué grupo pertenecen, la cantidad recibida de postulaciones al anuncio se va a ver afectada, por ejemplo si se trata de un grupo reducido, el anuncio recibirá menos postulaciones.

4. Postulantes

Sección dedicada al estudio de los potenciales postulantes. Potencialmente relevante para estudios futuros.

4.1. Cantidad de postulantes divididos por edad

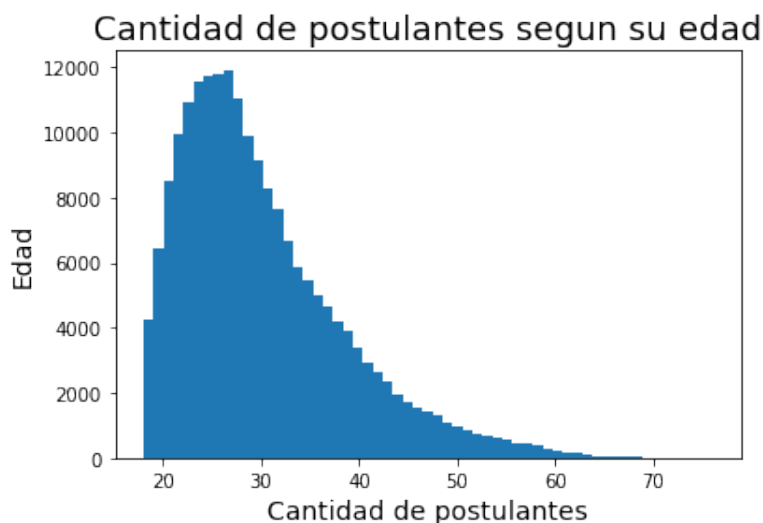


Figura 15: Cantidad de postulantes de cada edad

Si se quiere estudiar cuantos postulantes de un cierto rango de edad se presentan a un aviso, es importante entender la cantidad inicialmente disponible. En el gráfico propuesto se observa la cantidad de postulantes de cada edad disponibles, se observa que la mayoría de los postulantes está entre los 20 y 30 años. Se puede esperar que avisos que pidan de experiencia u otras cualidades que impliquen una edad mayor a los 30 años, reciban una menor cantidad de postulaciones que los que obvien estos requerimientos. Como consecuencia la posibilidad de que un postulante tenga una mayor edad aumenta a medida que se postule a avisos menos populares.

Un ejemplo para esto seria el gráfico **3.3.7 Pide experiencia?**

4.2. Cantidad de postulantes divididos por nivel educativo

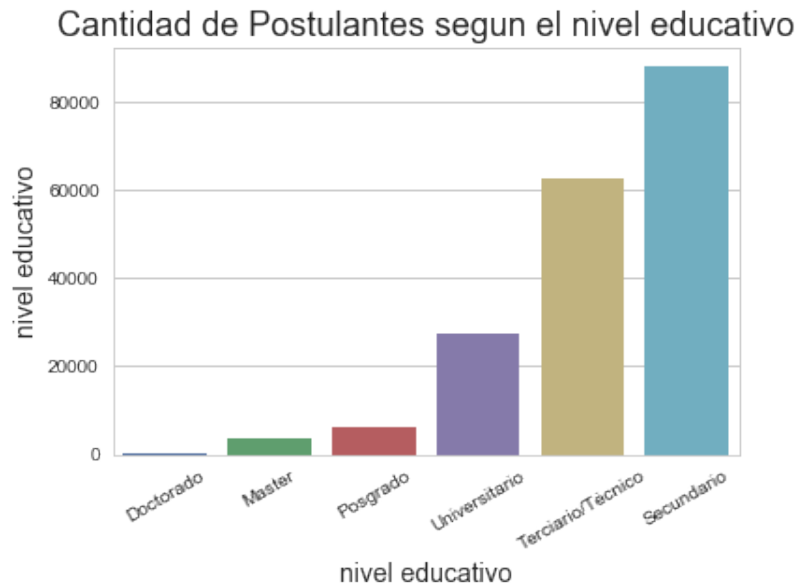


Figura 16: Cantidad de postulantes de cada nivel educativo

Un segundo agrupamiento posible es por nivel de educacion, en este caso se observa que la cantidad de postulantes en cada grupo disminuye drásticamente a medida que aumenta el nivel educativo. Esto implica 2 cosas, que anuncios que pidan un mayor nivel educativo reciban una menor cantidad de postulaciones y como consecuencia de esto, postulantes que se presenten a avisos menos populares probablemente posean una mejor preparacion. Ejemplo de esto se puede encontrar en **Promedio de postulaciones segun el tipo de trabajo**

5. Evolución Temporal

En esta sección se analizará la relación entre el tiempo y la cantidad de postulaciones y vistas a avisos. En particular, se estudiarán cómo varían esas cantidades en función de la hora del día, día de la semana y día del mes.

5.1. Postulaciones según hora del día

En la siguiente figura se muestra la cantidad de postulaciones registradas en función de la hora del día.



Figura 17: Cantidad de postulaciones a avisos de acuerdo con la hora del día

La mayor cantidad de postulaciones se registra por la mañana, donde el pico se observa alrededor de las 9 horas. A partir de ahí, el número disminuye casi constantemente cuánto más tarde se hace, hasta las 23 horas, cuando se nota una brusca caída. En los horarios de la madrugada se observa la mínima cantidad de postulaciones, en particular, a las 3 horas. Este patrón tiene sentido con una rutina de decisión y acción donde es frecuente que la decisión se tome un día y su acción sea lo primero que se realiza el día siguiente.

5.2. Vistas según hora del día

En la siguiente figura se muestra la cantidad de vistas registradas en función de la hora del día.

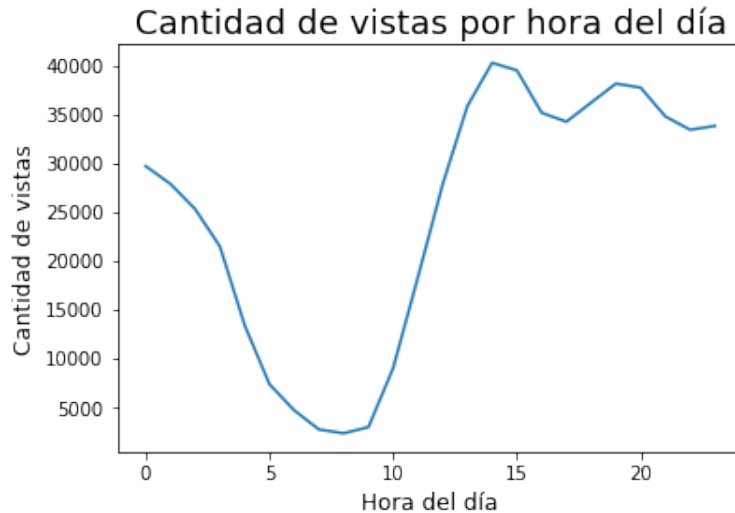


Figura 18: Cantidad de vistas a avisos de acuerdo con la hora del día

En este gráfico se observa que por la mañana se encuentra el mínimo de vistas mientras que por la tarde aumenta notoriamente y permanece alto hasta media madrugada. Este patrón tiene que ver con cuándo tiene una persona promedio tiempo libre para explorar las distintas posibilidades a empleo. Por las mañanas, es posible que una persona se encuentre durmiendo, en otro trabajo o estudiando. A eso de las 14 horas, durante el almuerzo, la gente suele tener tiempo libre, el cual puede utilizar para realizar dichas vistas. Luego disminuye de nuevo, probablemente por la ocupación o estudios, y vuelve a aumentar a eso de las 19 horas cuando se suele regresar a las casas. Finalmente, es común asimismo quedarse hasta tarde viendo opciones de empleo, lo cual justifica la cantidad de vistas por la madrugada.

5.3. Postulaciones según día de la semana

En la siguiente figura se muestra la cantidad de postulaciones registradas en función del día de la semana.

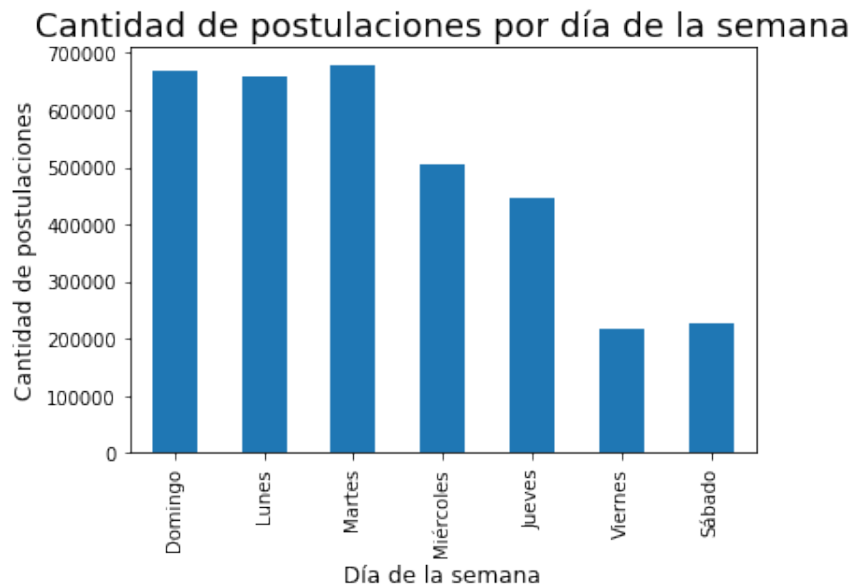


Figura 19: Cantidad de postulaciones a avisos de acuerdo con el día de la semana

La mayor cantidad de postulaciones se realizan los domingos, lunes y martes mientras que a medida que avanza la semana, disminuye rotundamente a un tercio de los valores.

5.4. Vistas según día de la semana

En la siguiente figura se muestra la cantidad de vistas registradas en función del día de la semana.

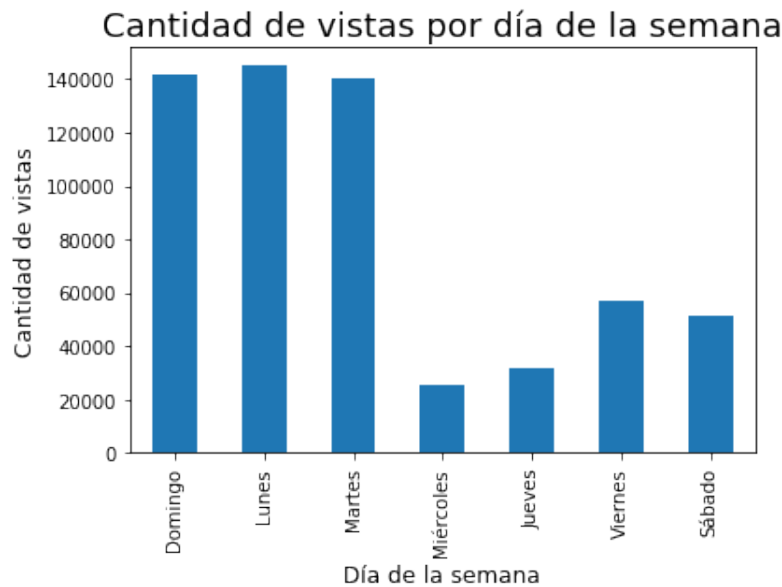


Figura 20: Cantidad de vistas a avisos de acuerdo con el día de la semana

Antes que nada, cabe destacar que esta información se obtiene de una única semana, del 23/2 al 1/3, y de este último día, se tiene información únicamente hasta las 4 de la madrugada. De nuevo se observa una gran cantidad de vistas el domingo, lunes y martes, mientras que el resto de los días registra menores cantidades.

5.5. Postulaciones según día del mes

Para este análisis, se estudió la cantidad de postulaciones de Enero y de Febrero por separado debido a la naturaleza de los datos.

5.5.1. Postulaciones según día del mes - Enero

En la siguiente figura se muestra la cantidad de postulaciones registradas en función del día del mes de Enero.

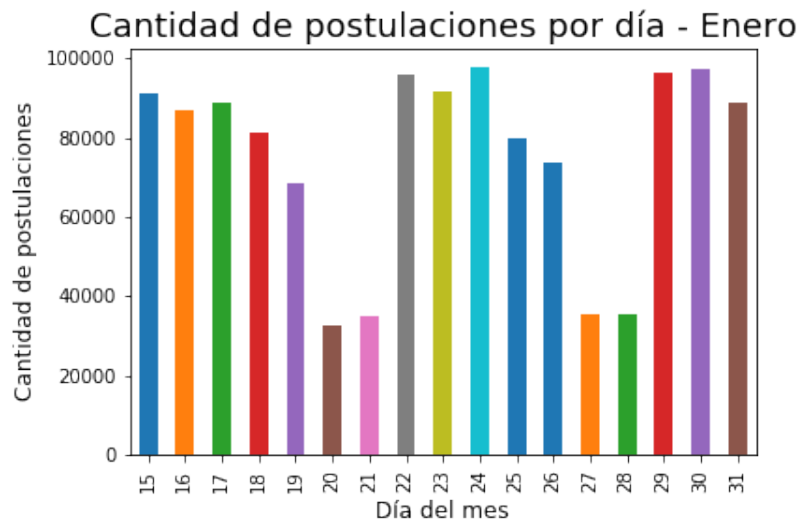


Figura 21: Cantidad de postulaciones a avisos de acuerdo con el día del mes: Enero

Esta data pertenece al mes de Enero, en el cual sólo se tiene información de la segunda mitad del mes. Los valles que se observan a los días 20, 21, 27 y 28 corresponden a fines de semana donde es más común que se descanse en vez de ocuparse de postularse a empleos. Los jueves y viernes anteriores a estos días, es decir los días 18, 19, 25, y 26 presentan una disminución de la cantidad de postulaciones, en preparación para el fin de semana. Más allá de esto, el resto de los días, que corresponden a lunes, martes y miércoles, muestran la mayor cantidad de postulaciones. Las personas suelen preocuparse sobre su futuro laboral principalmente durante el comienzo de la semana.

5.5.2. Postulaciones según día del mes - Febrero

En la siguiente figura se muestra la cantidad de postulaciones registradas en función del día del mes de Febrero.

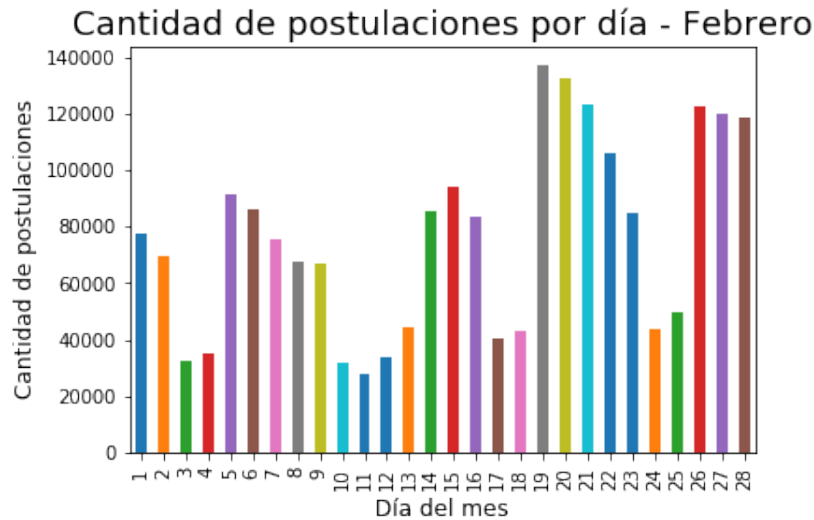


Figura 22: Cantidad de postulaciones a avisos de acuerdo con el día del mes: Febrero

Esta data pertenece al mes de Febrero, donde se tiene información sobre todo el mes. Al igual que en el gráfico anterior, se observan los valles en los fines de semana, con la adición de los días 12 y 13 de febrero que correspondieron a los feriados carnaval. Más allá de eso, presenta los mismos patrones del gráfico anterior: la cantidad de postulaciones es alta al comenzar la semana y disminuye a medida que se acerca el fin de la misma. Además, es interesante notar que la cantidad de postulaciones es mayor a fin de mes que a comienzo, fenómeno que puede ser que ocurra seguido, o puede tener que ver con mucha gente regresando de un viaje en la segunda quincena de Febrero.

5.6. Vistas según día del mes

En la siguiente figura se muestra la cantidad de vistas registradas en función del día del mes. Fue necesario filtrar una serie de registros que corresponden al jueves primero de marzo, donde la información alcanzaba únicamente el horario de las cuatro de la madrugada en vez de todo el día, distorsionando la interpretación de la información.

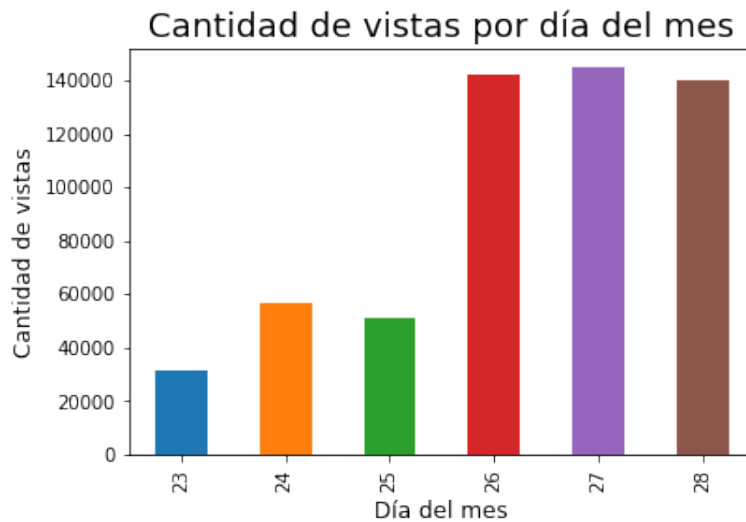


Figura 23: Cantidad de vistas a avisos de acuerdo con el día del mes

Esta información corresponde a las vistas realizadas desde el viernes 23/2 hasta el miércoles 28/2, es decir, seis días. Una vez más, se observa el mismo patrón: 26, 27 y 28 corresponden a lunes, martes y jueves, donde se registra la mayor cantidad de vistas, mientras que durante el fin de semana, la menor.

5.7. Conclusión sobre el aspecto temporal

De toda esta información se puede concluir que la mayor concentración de actividad, ya sea vista o postulación a aviso, ocurre los días domingo, lunes y martes, con la condición de que no se trate de un feriado. En cuanto a los horarios, se prefiere la mañana para postulaciones y las tardes/noches para vistas.

Esta data nos provee con conocimiento acerca del comportamiento de la gente que utiliza estos servicios. Con ella, la empresa puede tomar decisiones óptimas que permitan alcanzar un mayor número de visualizaciones por aviso, de modo de, consecuentemente, probablemente incrementar la cantidad de postulaciones a los mismos. Un ejemplo de esto sería planear en qué día y a qué hora es conveniente publicar los avisos, ya sea eligiendo un día de semana particular, o incluso un día de mes específico.

6. Nivel laboral y Nivel educativo.

En esta sección se analizará la relación entre el nivel del puesto laboral que se ofrece contra el nivel de educación que tenga el postulante a dicho puesto. Se pretendrá determinar si para puesto de alto niveles se postulan personas con mayor educación.

6.1. Postulaciones por nivel educativo.

En la siguiente figura se puede observar la cantidad de postulaciones por nivel educativo.

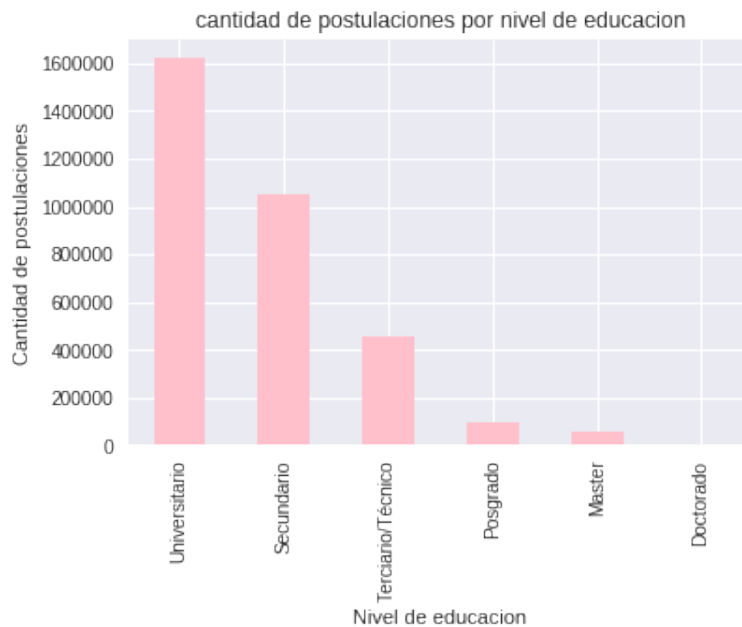


Figura 24: Cantidad de postulaciones según nivel educativo del postulante.

En primer instancia vemos que los principales postulantes tienen un nivel de estudio universitario. Nos preguntamos en qué estado está ese estudio:

Proporción de postulantes segun estado universitarios

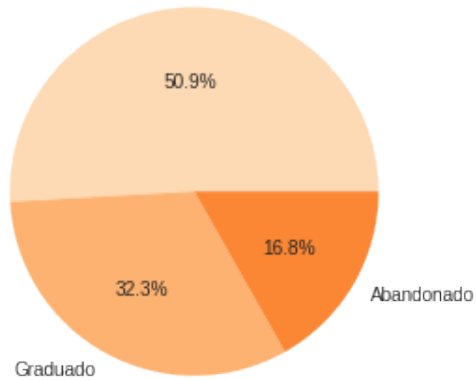


Figura 25: Proporción de postulaciones segun el estado de educación universitaria.

Vemos que la mayoría de los universitarios que se postulan son estudiantes.

Seguimos con los siguientes niveles de estudio:

Proporción de postulantes segun estado secundaria

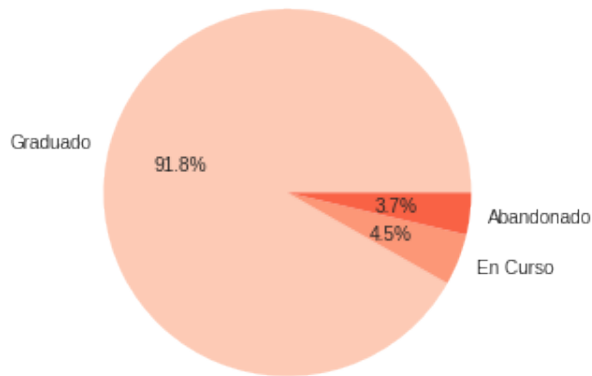


Figura 26: Proporción de postulaciones segun el estado de educación secundaria.

En este caso los postulantes en su mayoría han terminado la secundaria.

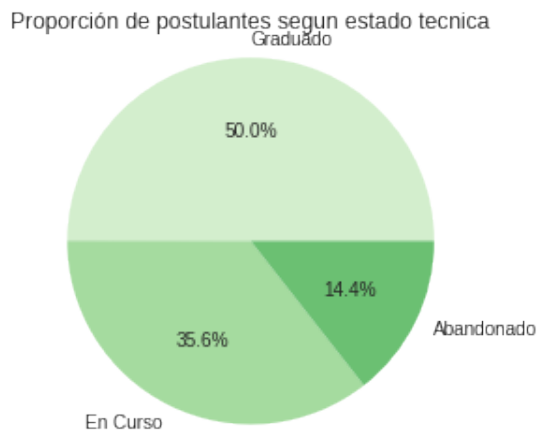


Figura 27: Proporción de postulaciones según el estado de educación técnica o terciaria.

Como se puede ver la mayoría de los postulantes son graduados para los que tienen un nivel terciario o técnico.

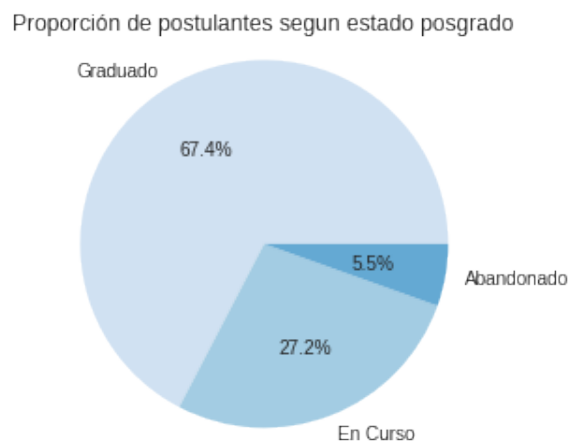


Figura 28: Proporción de postulaciones según el estado de educación de posgrado.

Se puede ver que en su mayoría son graduados.

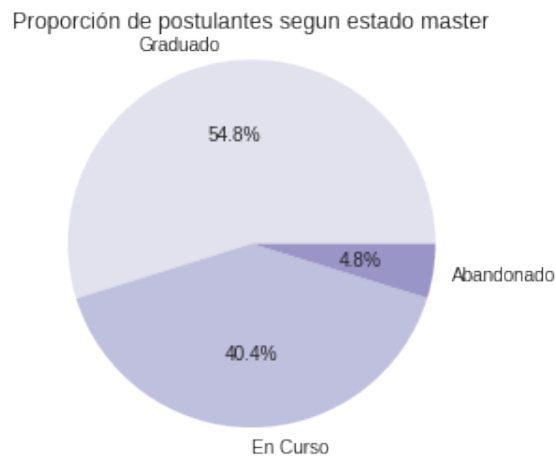


Figura 29: Proporción de postulaciones según el estado de educación de master.

En este caso la mayoría son graduados pero no hay tanta diferencia con los que se encuentran estudiando.



Figura 30: Proporción de postulaciones según el estado de educación de doctorado.

En este caso sucede lo mismo que quienes tienen un nivel de posgrado, es decir, no hay tanta diferencia entre la cantidad de postulantes graduados y en estudio.

6.2. Publicaciones por nivel laboral.

En esta sección se trata de ver que cantidad de publicaciones se dan por nivel laboral.

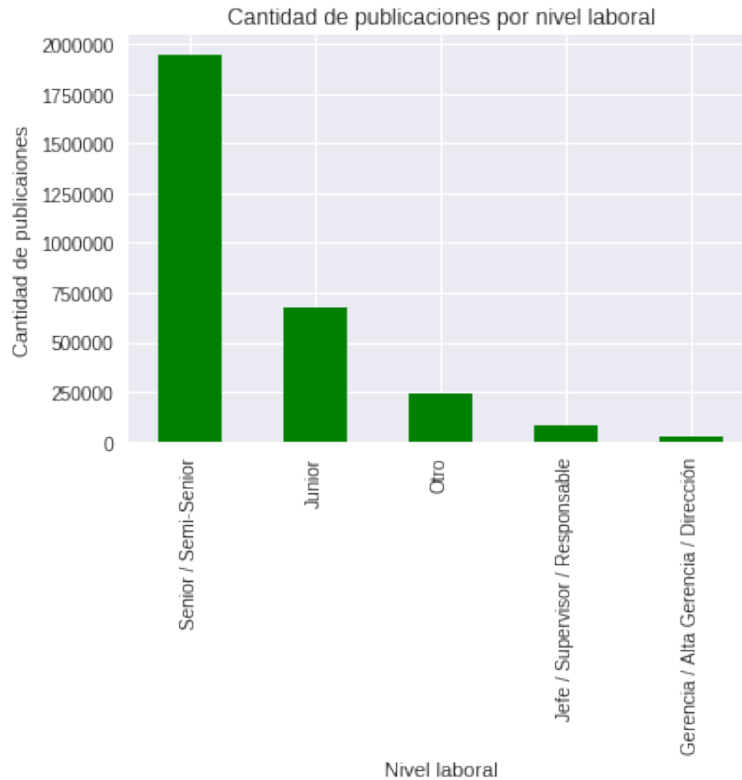


Figura 31: Cantidad de publicaciones por nivel laboral.

Vemos que hay mayor postulaciones a puestos de Senior o Semi-Senior

Nos preguntamos si las postulaciones a los cargos de mayor jerarquía se dan para personas con mayor nivel educativo. Es por eso que procedemos a agrupar por nivel laboral para comparar que nivel de educación tienen los postulantes a esos puestos:

6.2.1. Postulantes a Senior/Semisenior.

La siguiente figura muestra la proporción de postulantes para puestos de Senior y Semi-Senior segun el total de postulantes para cada nivel de educación.

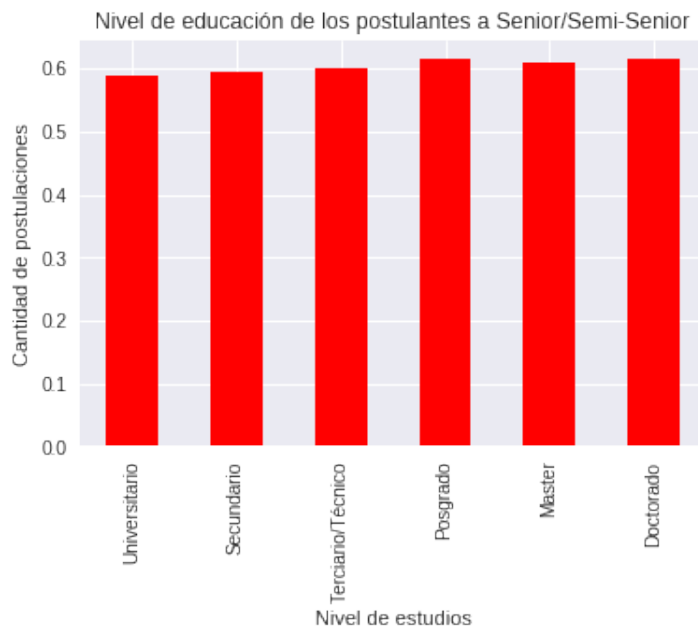


Figura 32: Áreas elegidas por postulantes con nivel de educación secundaria para puestos de Senior o Semi-Senior.

Como podemos ver la principal area a la cual los que tienen nivel secundario son esas 10, teniendo como principal a la Producción y Venta.

Analizamos ahora a que areas se postulan los universitarios, para ver si coinciden con los de secundario:

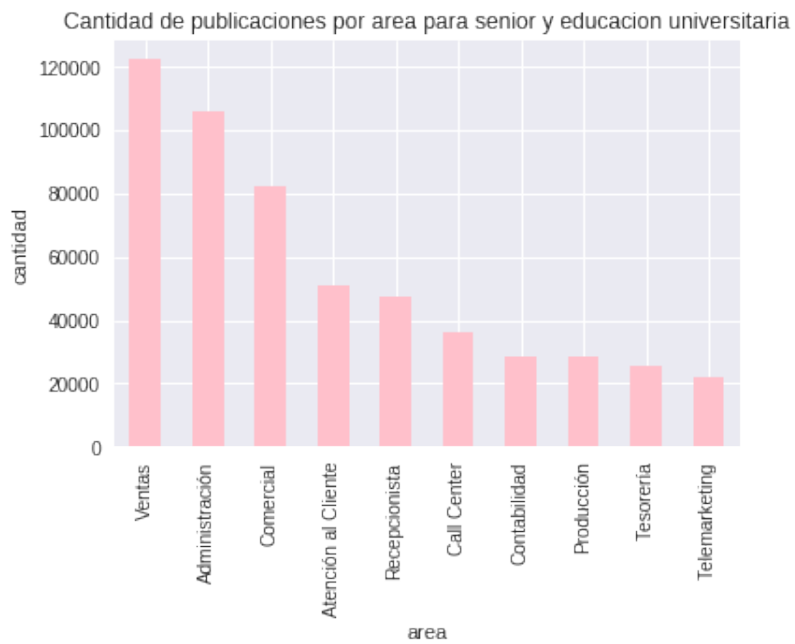


Figura 33: Áreas elegidas por postulantes con nivel de educación universitario para puestos de Senior o Semi-Senior.

Para los universitarios las principales areas de postulación son Ventas, administración y comercio.

Analizamos para postulantes de estudio terciario o técnico:

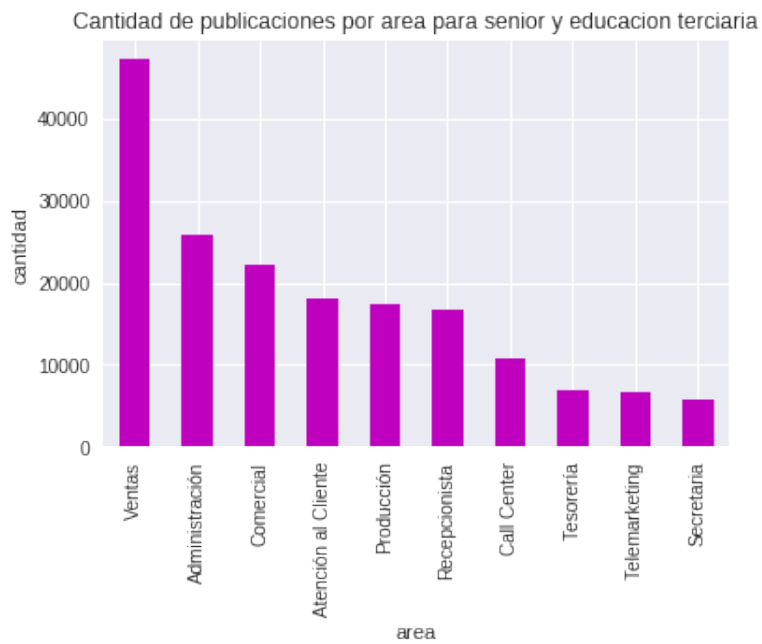


Figura 34: Áreas elegidas por postulantes con nivel de educación terciaria o técnica para puestos de Senior o Semi-Senior.

En este caso tambien hay una mayor recurrencia en ventas.
 Analizamos para postulantes de estudio de posgrado:

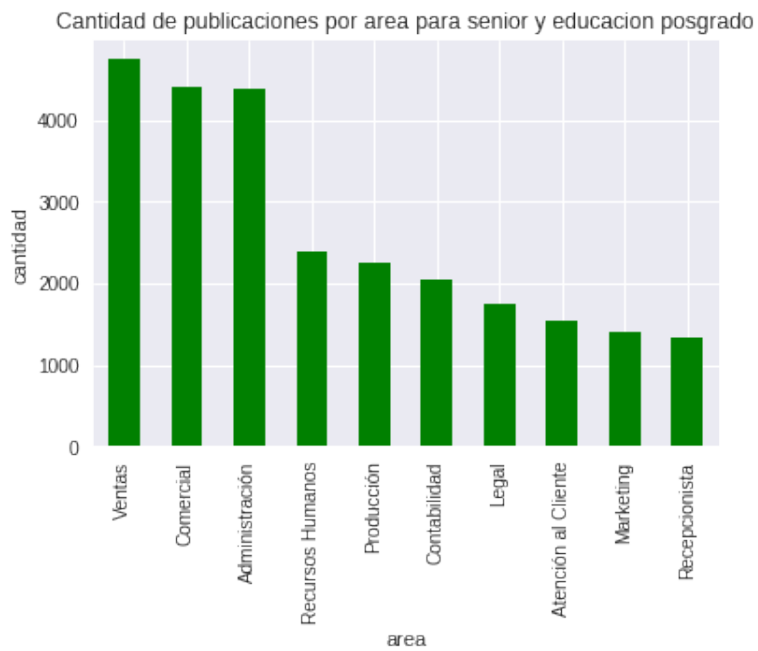


Figura 35: Áreas elegidas por postulantes con nivel de educación de posgrado para puestos de Senior o Semi-Senior.
Seguimos para quienes tienen un nivel de educación en master:

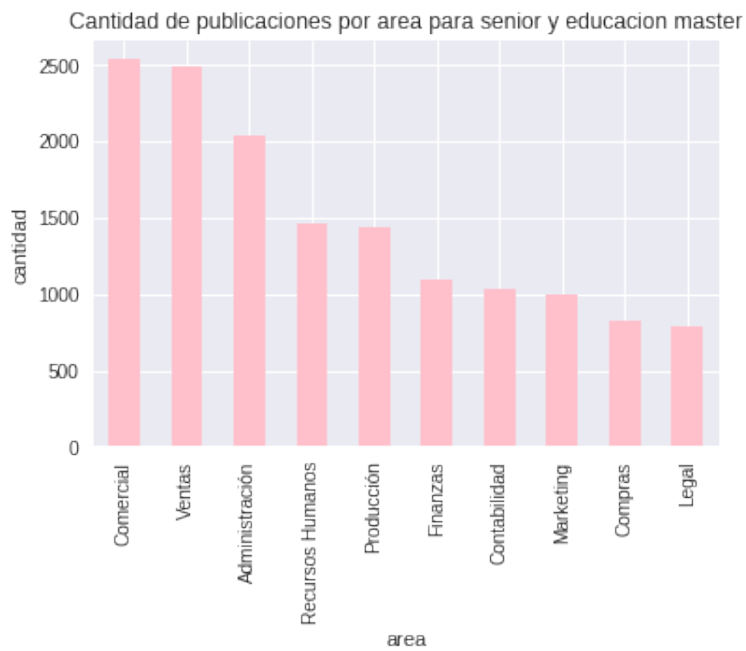
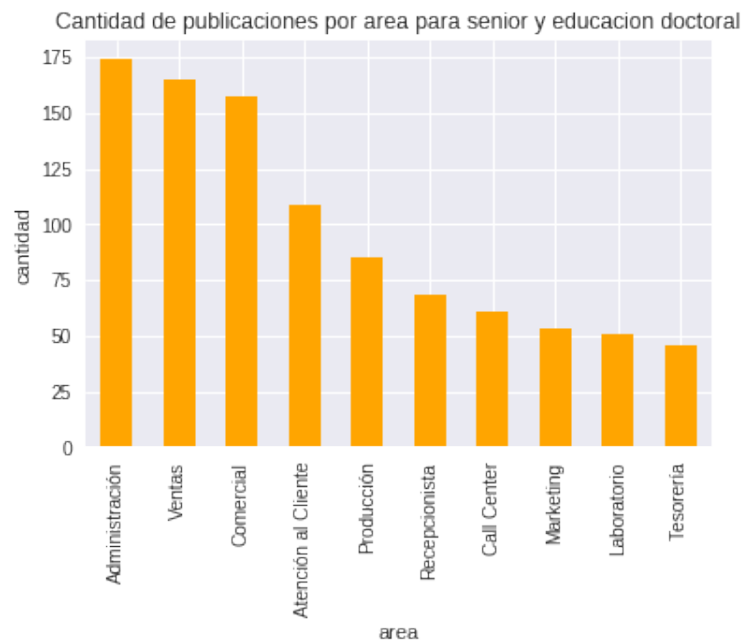


Figura 36: Áreas elegidas por postulantes con nivel de educación de master para puestos de Senior o Semi-Senior.

Segun el gráfico obtenido las áreas principales para quienes tienen un nivel de educación hasta master son Comercio, Ventas y Administración.

Por último vemos las áreas de preferencia para los que tienen nivel doctoral en su educación:



(1).png

Figura 37: Áreas elegidas por postulantes con nivel de educación doctoral para puestos de Senior o Semi-Senior.

Las áreas principales en este caso son administración, ventas y comercial.

6.2.2. Postulaciones para Junior.

En esta sección procedemos a ver que nivel de educación tienen los postulantes mas propensos a puestos de Junior.

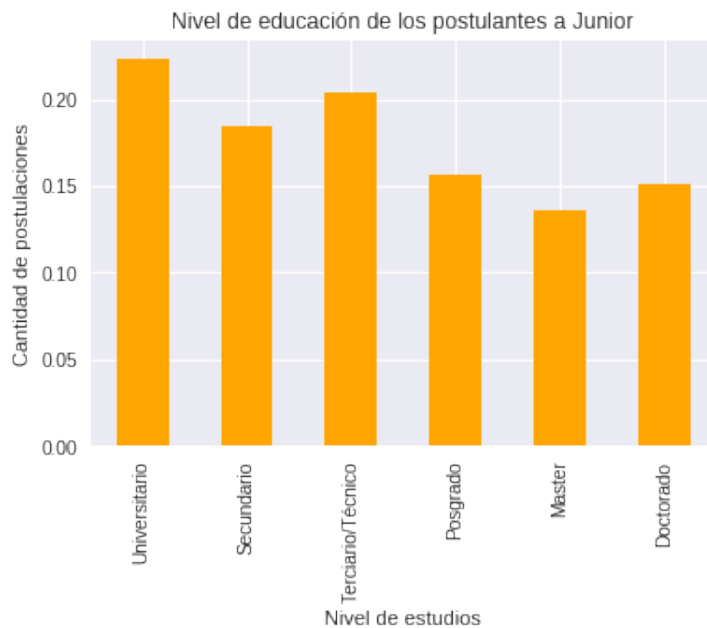


Figura 38: Proporción de postulaciones apuestos de Junior según nivel educativo.

En este caso, las proporciones son un poco mas diversas, aunque no con mucha diferencia. Siendo los de mayor concurrencia a postularse como Junior los universitarios y de estudio terciario.

6.2.3. Postulaciones para Jefe.

En esta sección procedemos a ver que nivel de educación tienen los postulantes mas propensos a puestos de Jefe.

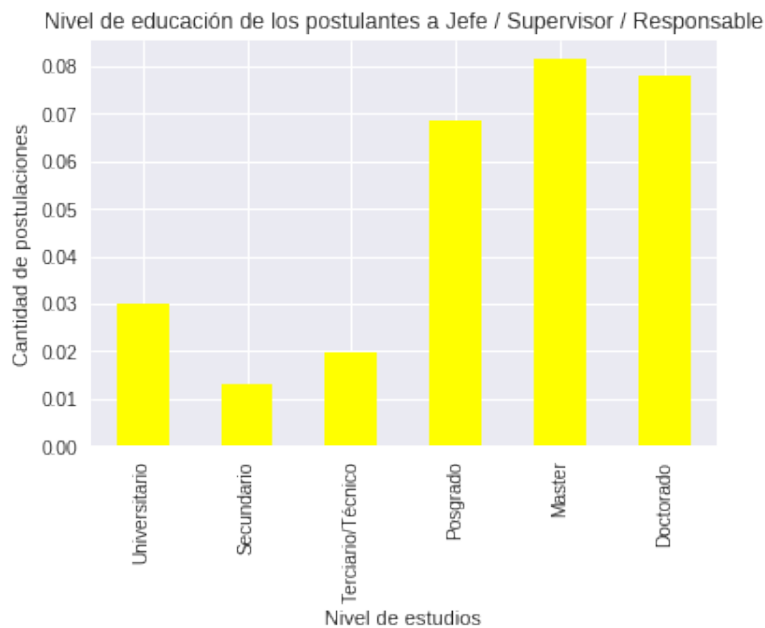


Figura 39: Proporción de postulaciones apostados de Jefe según nivel educativo.

En este caso vemos que los mas propensos a puestos de jefe son los que tienen niveles de educación de posgrado, doctorado y master.

6.2.4. Postulaciones para Gerencia / Alta Gerencia / Dirección.

En esta sección procedemos a ver que nivel de educación tienen los postulantes mas propensos a puestos de Gerencia / Alta Gerencia / Dirección.

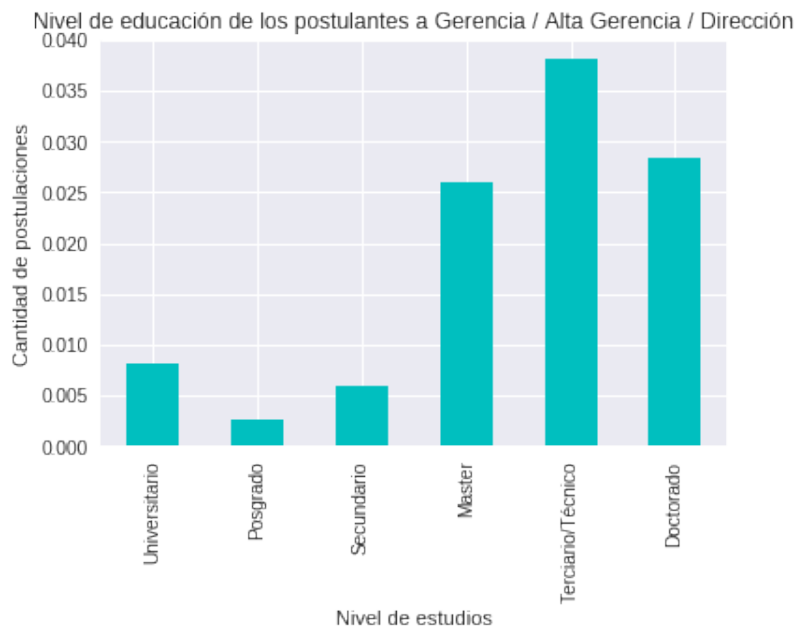


Figura 40: Proporción de postulaciones apuestos de Gerencia / Alta Gerencia / Dirección según nivel educativo.

En este caso vemos que los mas propensos a puestos de Gerencia / Alta Gerencia / Dirección son los que tienen niveles de educación de posgrado, doctorado y master, al igual que sucedía con los puestos de jefe.

6.3. Conclusión sobre nivel laboral y educacional.

Según pudimos observar existe una relación entre el nivel laboral solicitado y la educación de los postulantes. Es decir, que aquellos cuya formación académica es mayor, prefieren puestos de jefatura o gerencia, mientras que para los sectores de nivel laboral mas bajo presentan casi una misma cantidad de postulaciones respecto al total de postulantes según nivel académico.