Daniel Vázquez Lara / Samantha Pineda / Alhondra Gonzalez

Octubre 2024

Resumen

Análisis de base de datos de “Trabajo remoto y salud mental” para determinar si el trabajo remoto aporta cambios positivos para la empresa y sus trabajadores

PROYECTO FINAL DE DIPLOMADO ANÁLISIS Y CIENCIA DE DATOS

Universidad de Colima - Transición

1. **SELECCIÓN Y JUSTIFICACIÓN DE LA BASE DE DATOS**
   1. **Búsqueda de la Base de Datos**

Actualmente vivimos en un mundo donde lo único constante es el cambio y tenemos que adaptarnos.

Hace un par de años una pandemia vino a retarnos y a incluir de manera constante en nuestro vocabulario el término trabajo remoto (home office), sin embargo, este tipo de trabajo tiene más años existiendo.

El origen del teletrabajo se remonta a la década de 1970, en un contexto de crisis energética, cuando el alto costo del combustible y los largos desplazamientos al lugar de trabajo comenzaron a plantear serios desafíos económicos y ambientales. Jack Nilles, un ingeniero y excientífico de la NASA (frecuentemente citado como el padre del teletrabajo), comenzó a explorar la idea de trabajar a distancia como una solución a estos problemas, utilizando la tecnología de la información y las comunicaciones. Sin embargo, la expansión de Internet en la década de 1990 fue lo que realmente catalizó la adopción del teletrabajo, permitiendo a los empleados trabajar eficazmente desde cualquier lugar.

En los años 2000, el teletrabajo comenzó a ser visto no solo como una solución a problemas logísticos o económicos, sino también como una forma de mejorar la calidad de vida de los empleados, ofreciendo una mayor flexibilidad para equilibrar el trabajo y la vida personal.

La llegada de la década de 2020, marcada por la pandemia global de COVID-19, representó un punto de inflexión sin precedentes para el teletrabajo. Con las restricciones de movilidad y la necesidad de distanciamiento social, empresas de todo el mundo adoptaron rápidamente el trabajo remoto como medida de emergencia. Esta situación aceleró una transformación digital que, en muchos casos, se esperaba que tomara años en materializarse, demostrando la eficacia del teletrabajo incluso en circunstancias extremadamente desafiantes.

Este cambio en el tipo de trabajo retó a las empresas a cambiar su estructura organizacional, pero no sólo pensando en la funcionalidad y las ventajas hacia el negocio, sino también hacia el trabajador.

La cultura global está siendo transformada tomando como estandarte la calidad de vida de las personas porque ante tanto cambio no sólo adicionamos el término trabajo remoto, también incluimos el estrés laboral *(reacción que se produce ante exigencias y presiones laborales que no se ajustan a las capacidades y conocimientos de una persona).*

* 1. **Justificación de la Selección**
     1. **Descripción de la base de datos seleccionada**

La base de datos a analizar es “Trabajo remoto y salud mental” (Remote Work & Mental Health), la cual muestra las respuestas de 5,000 personas entrevistadas respecto a preguntas clave que pueden dar conclusiones clave respecto a cómo es percibido el trabajo remoto, su beneficio para las empresas y sobre todo para los trabajadores.

* + 1. **Justificación de la selección de la base de datos**

Seleccionamos esta base de datos porque nos parece interesante conocer las opiniones de personas alrededor del mundo respecto a una tendencia global, adicional, de que actualmente existen muchas empresas que han adoptado el trabajo remoto, sin embargo, también hay otras que después de haberlo tomado como base, no están logrando demostrar productivamente las ventajas esperadas, planteándose la posibilidad de regresar al antiguo estilo de trabajo en sitio o hacer una mezcla y tener un tipo de trabajo híbrido.

* + 1. **Contexto del problema**

Analizar si el trabajo remoto aporta cambios positivos para la empresa y para los trabajadores comparando las siguientes variables

|  |  |  |
| --- | --- | --- |
| * Edad | * Puesto | * Tipo de industria |
| * Tipo de trabajo | * Niveles de estrés | * Afectación a la salud |
| * Acceso a cuidado de la salud | * Productividad | * Satisfacción |
| * Actividad física | * Calidad de sueño |  |

Para alcanzar el objetivo anterior, se analizaron respuestas de 5,000 personas encuestadas de diferentes regiones (África, Asia, Europa, Oceanía, América del Norte y América del Sur).

* + 1. **Relevancia de los datos**

Los datos son el activo más importante en la toma de decisiones, sin embargo, no basta con tener datos, hay que saber analizarlos y utilizarlos para poder optimizarlos.

* + 1. **Objetivos esperados**

Definir:

* Cuál es la región (de las mencionadas) que realiza más trabajo remoto
* Cuál es la industria que puede realizar trabajo remoto
* Cuál es el puesto que es más frecuente que realice trabajo remoto
* Si el trabajo remoto incrementa los niveles de estrés
* Si el trabajo remoto genera alguna afectación a la salud mental
* Si las personas que tienen alguna afectación mental tienen acceso al cuidado de la salud
* Si el trabajo remoto impide la realización de actividad física
* Si el trabajo remoto interfiere con la calidad del sueño de los trabajadores
* Si las personas se sienten satisfechas con el desarrollo del trabajo remoto
* Si el trabajo remoto ha incrementado la productividad de los trabajadores

1. **PREPARACIÓN Y LIMPIEZA DE DATOS**
   1. **Descripción de la Base de Datos**
2. **Número de registros y variables/campos**

La base de datos seleccionada (archivo CSV) cuenta con 5,000 registros con 20 columnas (categorías).

Para fines de análisis de este proyecto, las columnas (categorías) utilizadas, fueron cambiando, dependiendo el requerimiento.

Las columnas (categorías) que no fueron utilizadas en el análisis fueron:

* Gender – no requerido
* Work\_Life\_Balance\_Rating – falta de información
* Social\_Isolation\_Rating – falta de información
* Company\_Support\_for\_Remote\_Work – falta de información

1. **Tipos de datos (numéricos, categóricos, fechas, etc.)**

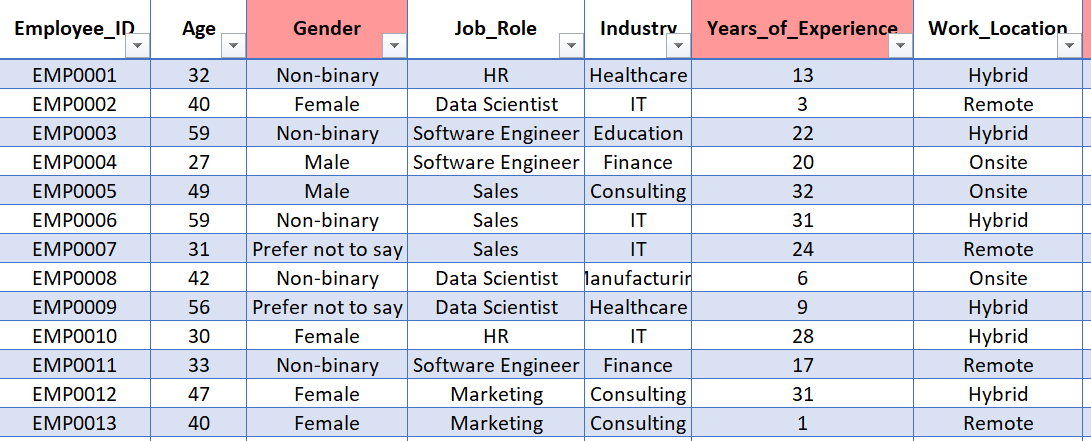
Los datos por analizar en este proyecto son numéricos y categóricos dependiendo la categoría.

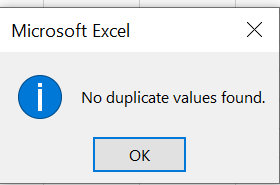
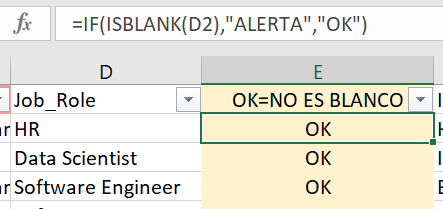
| **CATEGORÍA** | **TIPO DE DATO** |  | **CATEGORÍA** | **TIPO DE DATO** |
| --- | --- | --- | --- | --- |
| Age | Numérico discreto |  | Job\_Role | Categórico |
| Industry | Categórico |  | Years\_of\_Experience | Numérico discreto |
| Work\_Location | Categórico |  | Hours\_Worked\_Per\_Week | Numérico discreto |
| Number\_of\_Virtual\_Meetings | Numérico discreto |  | Stress\_Level | Categórico |
| Mental\_Health\_Condition | Categórico |  | Access\_to\_Mental\_Health\_Resources | Categórico |
| Productivity\_Change | Categórico |  | Satisfaction\_with\_Remote\_Work | Categórico |
| Physical\_Activity | Categórico |  | Sleep\_Quality | Categórico |
| Region | Categórico |  |  |  |

* 1. **Limpieza de Datos**

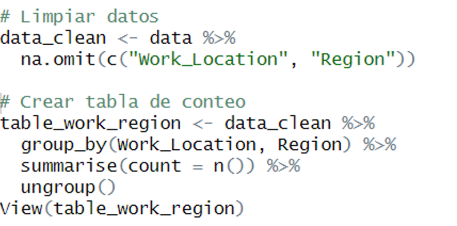
Para realizar la limpieza de datos se realizó de dos maneras:

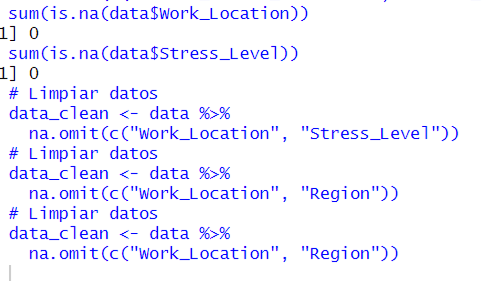
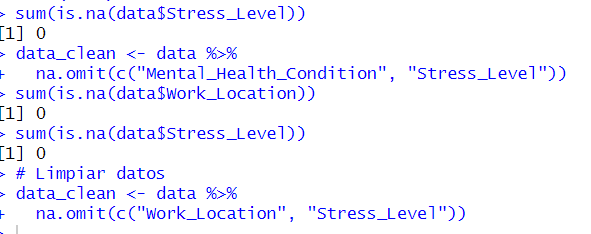
* **Excel. No se detecta necesidad de realizar limpieza.**
  + Se seleccionan todos los datos y se activa el formato tabla.
  + Se tomó como base el número de empleado y se revisó que no estuviera duplicado.
  + Se revisó por medio de filtros que no se tuvieran datos en blanco en cada una de las columnas (categorías).





* **R. No se detecta necesidad de realizar limpieza.**
  + Se analiza mediante R por medio de códigos de limpieza, arrojando que no hay evidencia de datos nulos o faltantes (algunos ejemplos se muestran en las siguientes imágenes).

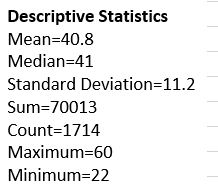


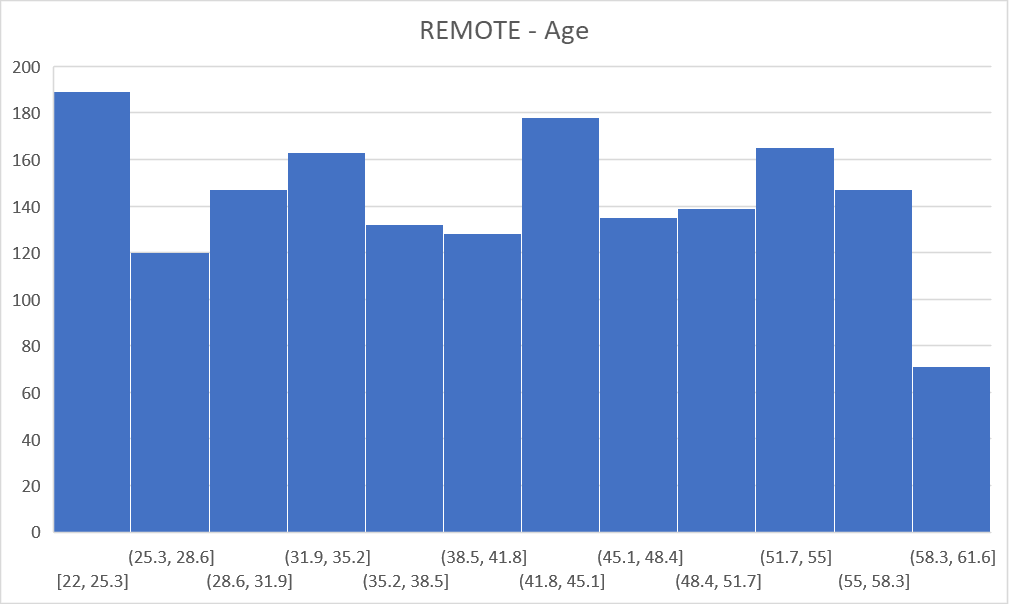
****

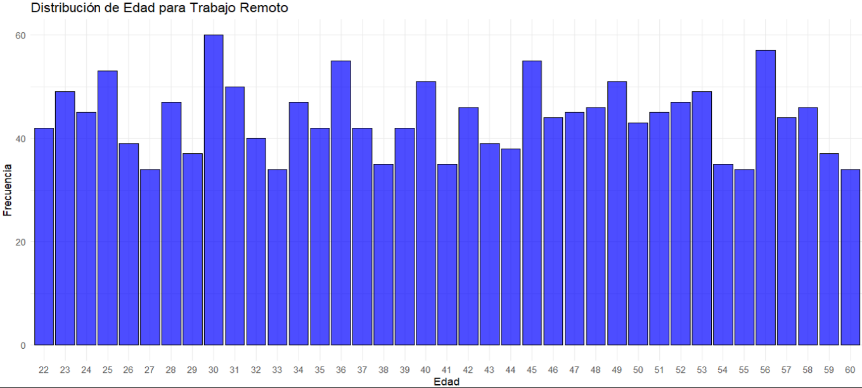
1. **ANÁLISIS EXPLORATORIO DE DATOS (EDA)**
   1. **Análisis Descriptivo (estadística descriptiva y distribución de variables)**

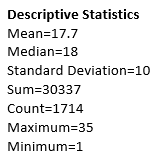
Se realizó el análisis descriptivo utilizando el software Excel para el cálculo de la estadística básica.

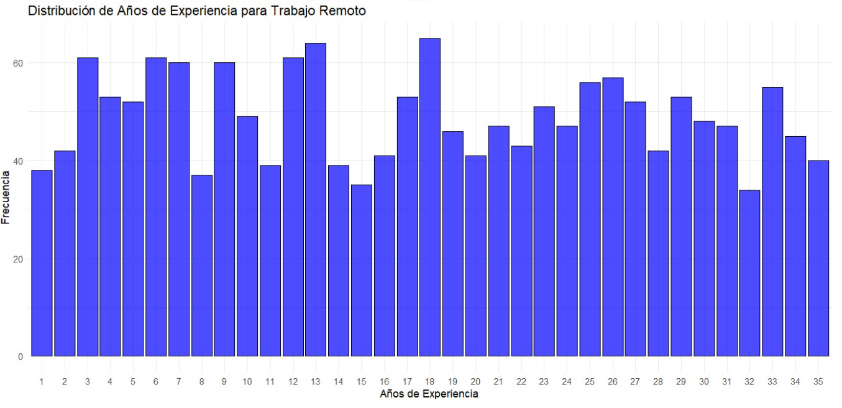
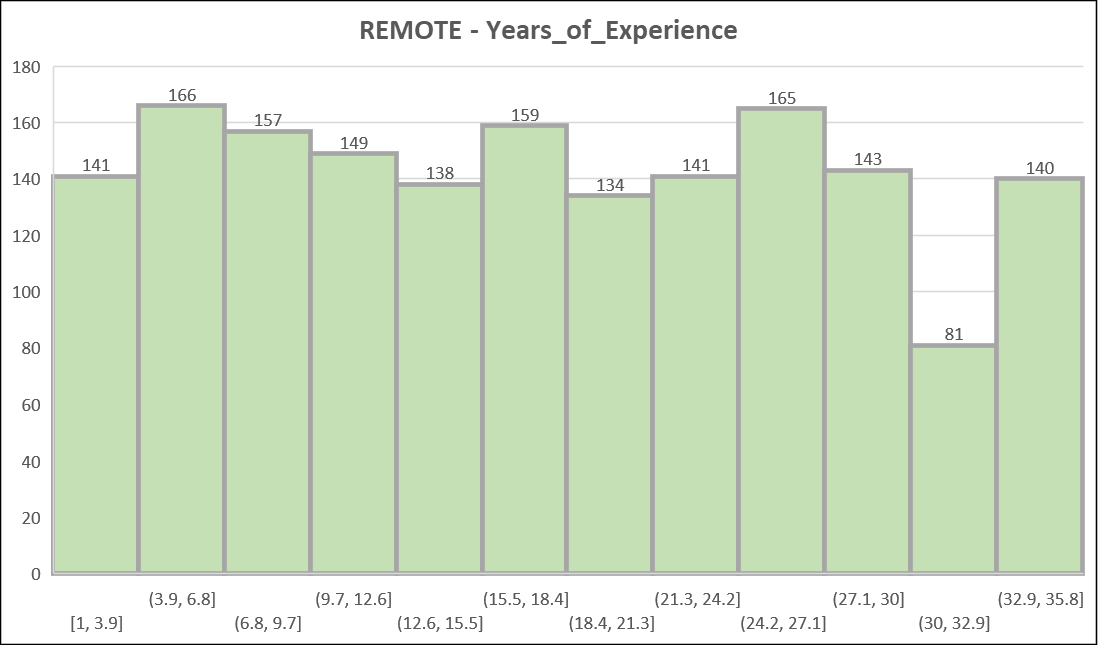
Para la sección de representación de las distribuciones de variables, se utilizó Excel y R para comparar el tipo de gráficos obtenidos.



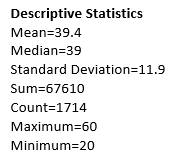


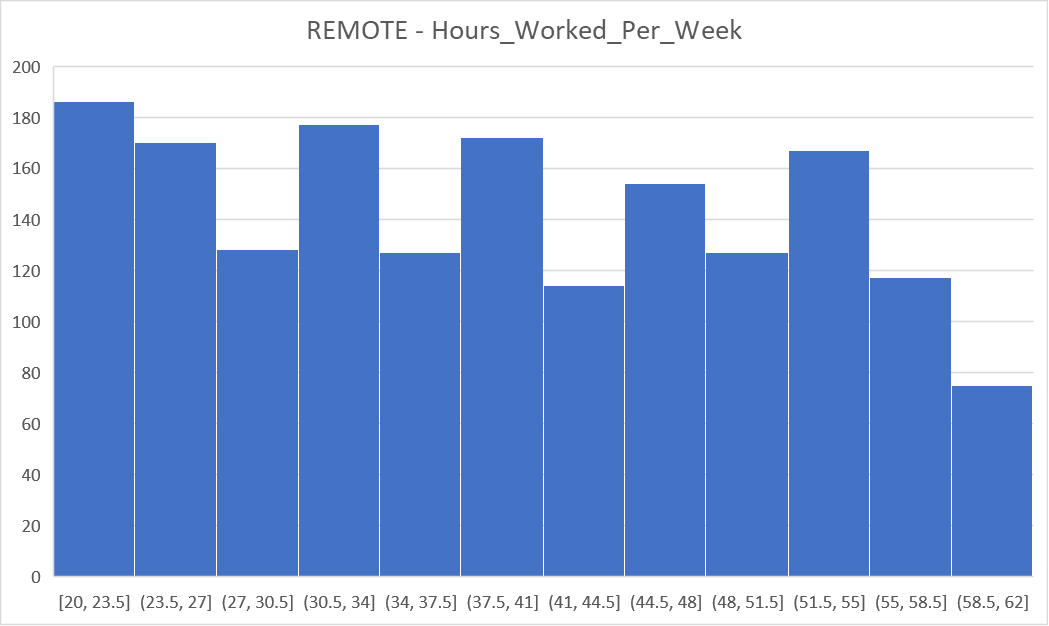
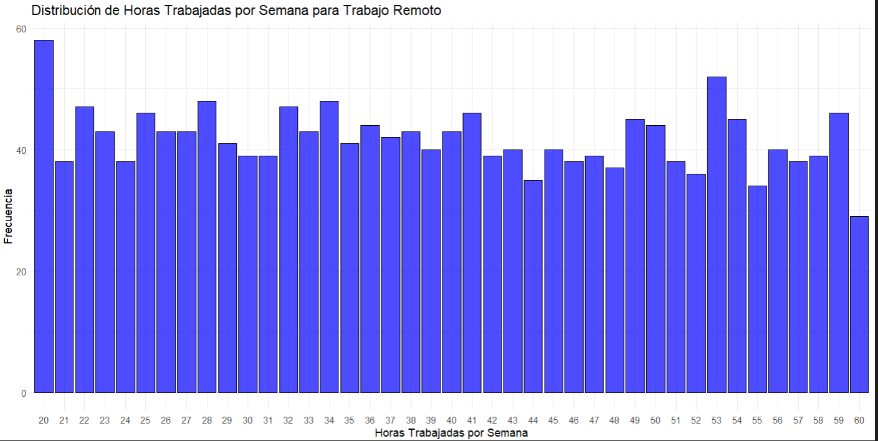
Interpretación. La edad promedio de las personas encuestadas que realizan trabajo remoto es de 40 años. De las personas entrevistadas que realizan trabajo remoto, 60 tienen 30 años (siendo el 4% del total de las personas que realizan trabajo remoto). De acuerdo con la desviación estándar, los datos se encuentran lejos de la media, lo que concluye que tienen mucha variabilidad.



****

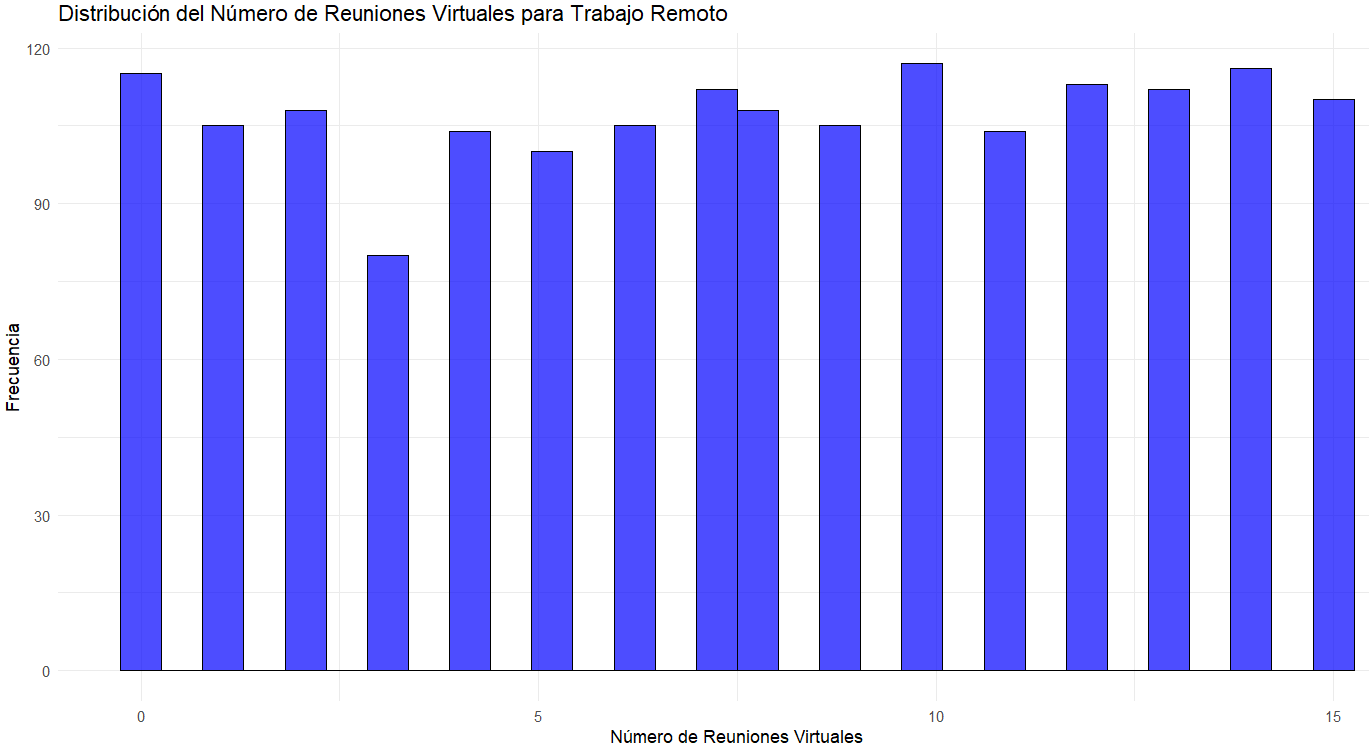
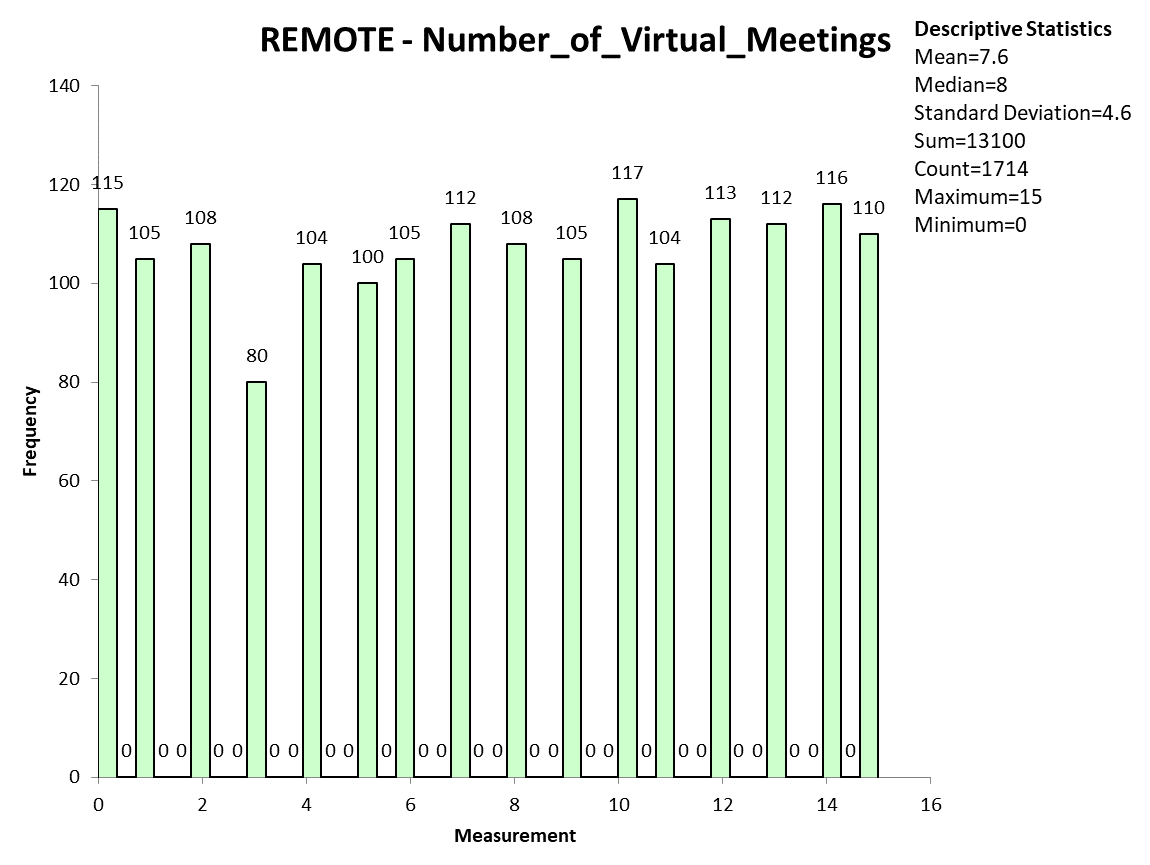
Interpretación. Los años de experiencia promedio de las personas encuestadas que realizan trabajo remoto es de 17.7 años. De las personas entrevistadas que realizan trabajo remoto, 65 tienen 18 años de experiencia (siendo el 4% del total de las personas que realizan trabajo remoto). De acuerdo con la desviación estándar, los datos se encuentran lejos de la media, lo que concluye que tienen mucha variabilidad.

****



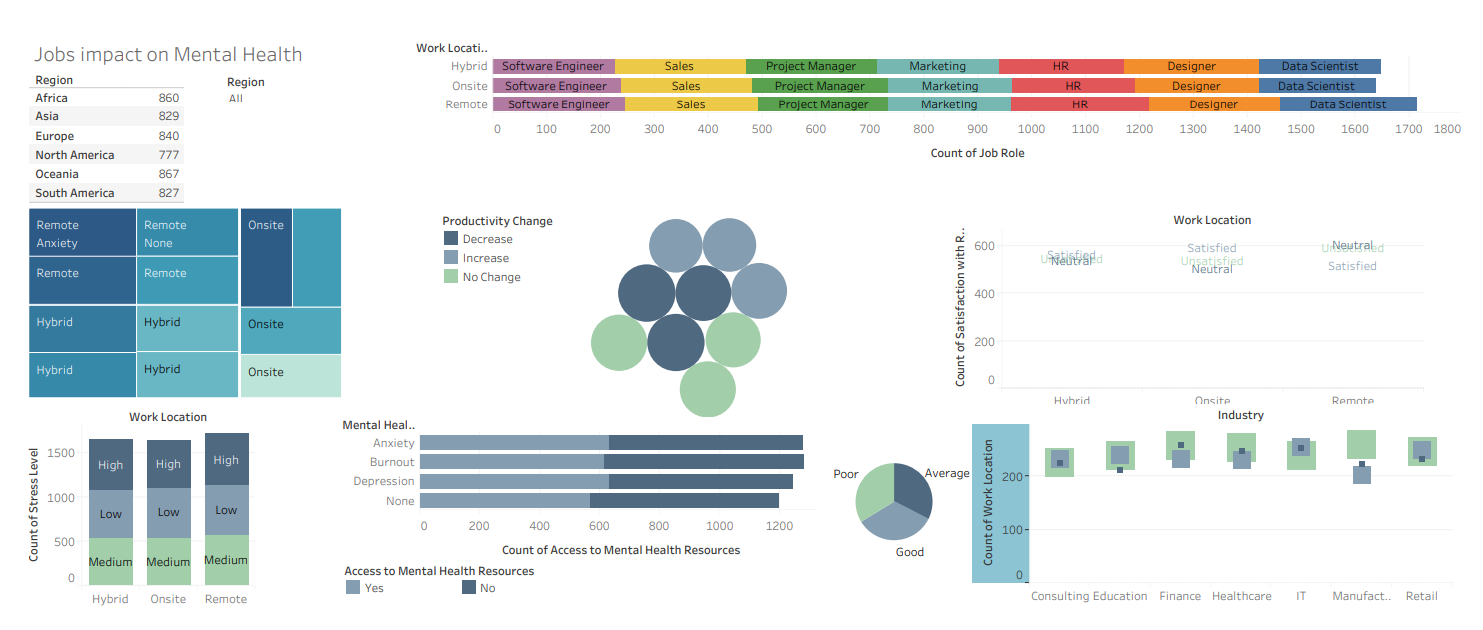
Interpretación. Las horas de trabajo semanales promedio de las personas encuestadas que realizan trabajo remoto es de 39.45 horas. De las personas entrevistadas que realizan trabajo remoto, 58 trabajan 20 horas a la semana (siendo el 3% del total de las personas que realizan trabajo remoto). De acuerdo con la desviación estándar, los datos se encuentran lejos de la media, lo que concluye que tienen mucha variabilidad.



****

Interpretación**.** El número de reuniones virtuales promedio en las que participan las personas encuestadas que realizan trabajo remoto es de 7.64 horas. De las personas entrevistadas que realizan trabajo remoto, 117 participan en 10 reuniones virtuales (siendo el 7% del total de las personas que realizan trabajo remoto). De acuerdo con la desviación estándar, los datos se encuentran lejos de la media, lo que concluye que tienen mucha variabilidad.

1. **VISUALIZACIÓN DE DATOS**
   1. **Herramienta de visualización**

****Para la visualización de datos utilizamos Tableau representando las principales gráficas.

[*https://public.tableau.com/app/profile/daniel.v.zquez.lara/viz/Dashboardproyectofinal1/Dashboard1*](https://public.tableau.com/app/profile/daniel.v.zquez.lara/viz/Dashboardproyectofinal1/Dashboard1)

1. **INTERPRETACIÓN Y CONCLUSIONES**
   1. **Resultados del Análisis**

* ¿Cuál es la región (de las mencionadas) que realiza más trabajo remoto?

Las personas que más trabajo remoto realizan se encuentran en África, seguido de Asia y Oceanía

* ¿Cuál es la industria que puede realizar trabajo remoto?

El tipo de industria que más realiza trabajo remoto es Manufacturing, seguido de Finance y Healthcare

* ¿Cuál es el puesto que es más frecuente que realice trabajo remoto?

La posición que más trabajo remoto realiza es Data Scientist seguido de Software Engineer

* ¿El trabajo remoto incrementa los niveles de estrés?

El trabajo remote genera mayores niveles de estrés (con 1% mayor). Los trabajos tipo en sitio e híbrido generan menores niveles de estrés (siendo ligeramente mejor en sitio)

* ¿El trabajo remoto genera alguna afectación a la salud mental?

El tipo de trabajo no está correlacionado con alguna afectación a la salud

* ¿Las personas que tienen alguna afectación mental tienen acceso al cuidado de la salud?

La mayoría de las personas que tienen alguna afectación a la salud no tienen acceso a atención médica

* ¿El trabajo remoto impide la realización de actividad física?

Al analizar todas las regiones, las personas que realizan trabajo remoto realizan más ejercicio diario que las que hacen trabajo híbrido o en sitio.

* ¿El trabajo remoto interfiere con la calidad del sueño de los trabajadores?

Considerando todas las regiones, la calidad de sueño no se ve afectada por el trabajo remoto (excepto en África en donde el trabajo remoto genera mejor calidad de sueño y Oceanía que el trabajo remoto general mala calidad de sueño). Detallando los datos, se puede concluir que el tipo de trabajo no afecta con la calidad del sueño.

* ¿Las personas se sienten satisfechas con el desarrollo del trabajo remoto?

La mayoría de las personas entrevistadas no están satisfechas con realizar trabajo remoto, a excepción de Europa.

* ¿El trabajo remoto ha incrementado la productividad de los trabajadores?

Las personas entrevistadas, consideran que al realizar actividades remotas su productividad ha disminuido.

* 1. **Conclusiones**

Después de realizar el análisis de los datos considerando el 80% de las variables presentadas se concluye que los datos no son completamente significativos puesto que no se tuvo una muestra homogénea de entrevistados por región, por industria y/o puesto.

Con los datos analizados se concluye que el trabajo remoto no es la mejor opción para incrementar los niveles de producción del personal ni mejorar su calidad de vida, sin embargo, no se tiene evidencia suficiente para determinar que el trabajo remoto afecta la salud mental de los trabajadores.

La finalidad del estudio es buena, pero se tienen áreas de oportunidad en la determinación de los datos a analizar, se recomienda se repita el estudio considerando los puntos mencionados.