

UNIVERSIDAD PERUANA UNIÓN
FACULTAD DE INGENIERÍA Y ARQUITECTURA
Escuela Profesional de Ingeniería Ambiental



**Modelamiento de accidentes de trabajo por actividad
económica mediante técnica de machine learning**

Tesis para obtener el Título Profesional de Ingeniero Ambiental

Autor:

Katherine Jasmin Enciso Carhuamaca
Raul Luis Huaytalla Ramirez
Nadyely Del Rosario Lino Valenzuela

Asesor:

Mg. Jackson Edgardo Perez Carpio

Ñaña, 07 de noviembre del 2023

DECLARACIÓN JURADA DE ORIGINALIDAD DE TESIS

Yo Jackson Edgardo Perez Carpio, docente de la Facultad de Ingeniería y Arquitectura, Escuela Profesional de Ingeniería Ambiental, de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: **“MODELAMIENTO DE ACCIDENTES DE TRABAJO POR ACTIVIDAD ECONÓMICA MEDIANTE TÉCNICA DE MACHINE LEARNING”** de los autores Katherine Jasmin Enciso Carhuamaca, Raul Luis Huaytalla Ramirez y Nadyely Del Rosario Lino Valenzuela tiene un índice de similitud de 16. % verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de Lima., a los 5 días del mes de diciembre del año 2023.



Jackson Edgardo Perez Carpio

ACTA DE SUSTENTACIÓN DE TESIS

En Lima, Ñaña, Villa Unión, a los 07 días día(s) del mes de noviembre del año 2023 siendo las 08:30 horas, se reunieron en modalidad virtual u online sincrónica, bajo la dirección del Señor Presidente del jurado: **Mg. Milda Amparo Cruz Huaranga**, el secretario: **Mg. Joel Hugo Fernández Rojas**, y los demás miembros: **Mg. Iliana Del Carmen Gutiérrez Rodríguez** y el Ing. **Orlando Alan Poma Porras**, y el asesor, **Mg. Jackson Edgardo Perez Carpio** con el propósito de administrar el acto académico de sustentación de la tesis titulada: "Modelamiento de accidentes de trabajo por actividad económica mediante técnica de machine learning"

de el(los)/la(las) bachiller(es): a) **KATHERINE JASMIN ENCISO CARHUAMACA**

.....b) **RAUL LUIS HUAYTALLA RAMIREZ**

.....c) **NADYELY DEL ROSARIO LINO VALENZUELA**

, conducente a la obtención del título profesional de **INGENIERO AMBIENTAL**
(Nombre del Título profesional)

El Presidente inició el acto académico de sustentación invitando al (los)/a(la)(las) candidato(a)/s hacer uso del tiempo determinado para su exposición. Concluida la exposición, el Presidente invitó a los demás miembros del jurado a efectuar las preguntas, y aclaraciones pertinentes, las cuales fueron absueltas por el(los)/la(las) candidato(a)/s. Luego, se produjo un receso para las deliberaciones y la emisión del dictamen del jurado.

Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Candidato (a): **KATHERINE JASMIN ENCISO CARHUAMACA**

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	18	A-	Muy Bueno	Sobresaliente

Candidato (b): **RAUL LUIS HUAYTALLA RAMIREZ**

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	18	A-	Muy Bueno	Sobresaliente

Candidato (c): **NADYELY DEL ROSARIO LINO VALENZUELA**

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	18	A-	Muy Bueno	Sobresaliente

(*) Ver parte posterior

Finalmente, el Presidente del jurado invitó al(los)/a(la)(las) candidato(a)/s a ponerse de pie, para recibir la evaluación final y concluir el acto académico de sustentación procediéndose a registrar las firmas respectivas.

Presidente
Mg. Milda Amparo
Cruz Huaranga



/Secretario
Mg. Joel Hugo
Fernández Rojas

Asesor
Mg. Jackson Edgardo
Perez Carpio

Miembro
Mg. Iliana Del Carmen
Gutiérrez Rodríguez

Miembro
Ing. Orlando Alan
Poma Porras

Candidato/a (a)
Katherine Jasmin

Candidato/a (b)
Raul Luis

Candidato/a (c)
Nadyely Del Rosario

Modelamiento de accidentes de trabajo por actividad económica mediante técnica de machine learning

Workplace accident modeling by economic activity using machine learning techniques

Resumen

En este artículo se expone la predicción de accidentes de trabajo dentro del sector manufacturero a través del modelo machine learning, debido a una mayor tendencia de accidentes en este sector. Para esto se muestreó y se recopiló datos de accidentes de trabajo en los 24 departamentos del Perú, a fin de predecir accidentes de trabajo futuros que podríamos tener. La metodología establece rangos de valor proyectados, donde se empleó logaritmos y calculados de modelos, para optimizar los riesgos de los accidentes de trabajo ocurridos en todo el país. Acorde al estudio realizado en el periodo del 2016 al 2021, se evidencia una tendencia al incremento en accidentes, es decir que cada año habrá un incremento de 481.97 accidentes. Los resultados en ese periodo de 6 años, se clasifican en accidentes de trabajo como lesiones, invalidez o muerte, aquellos que van a ocurrir y con mayor frecuencia por la naturaleza de sus actividades.

Palabras clave

Machine Learning; Accidente de trabajo; Predicción; Actividad Económica.

Abstract

In this article, the prediction of accidents at work within the manufacturing sector is exposed through the machine learning model, due to a greater tendency of accidents in this sector. For this, work accident data was sampled and collected in the 24 departments of Peru, in order to predict future work accidents that we could have. The methodology establishes projected value ranges, where logarithms and calculated models were used, to optimize the risks of work accidents that occurred throughout the country. According to the study carried out in the period from 2016 to 2021, there is a trend towards an increase in accidents that is to say that each year there will be an increase of 481.97 accidents. The results in this 6-year period are classified as work accidents such as injuries, disability or death, those that will occur and more frequently due to the nature of their activities.

Keywords

Machine learning; Work accident; Prediction; Economic activity.

Introducción

La exposición a 19 factores de riesgo ocupacionales específicos evidenció que acarrearón en todo el mundo casi 2 millones de muertes entre 2010 y 2016, el 82% de las cuales se debieron a enfermedades ocupacionales, la principal causa de muertes son las enfermedades cardiovasculares. Para las organizaciones empresariales, esto puede tener impactos en la productividad, incluidas las pérdidas debidas al deterioro permanente y los costos de rotación del personal. Además, es de estimar que los accidentes laborales y las enfermedades ocupacionales provocan la pérdida del 5,4% del producto interno bruto mundial anualmente, lo que implica en el avance económico (1,2).

De acuerdo al informe difundido por la Organización Internacional del Trabajo (OIT), los trabajadores de la construcción en los países desarrollados tienen una probabilidad casi cuatro veces mayor de verse involucrados en un accidente de muerte que los trabajadores de otros sectores industriales, mientras que los colaboradores del sector construcción en países en vías de desarrollo tienen casi seis veces más posibilidades que otros trabajadores (3,4) . Según la Organización Internacional de Trabajo (OIT), diariamente se producen 6.300 muertes a causa de accidentes o enfermedades relacionadas con el ámbito laboral remunerado, lo que suma más de 2,3 millones de fallecimientos al año en nivel global. Se proyecta que cada año suceden más de 317 millones de accidentes de trabajo, es decir que cada 15 segundos 153 personas que trabajan sufren un accidente laboral, siendo muchos de éstos lesiones con baja laboral (5) .

Según el (Organización Mundial de la Salud, 2017) señala que 2 millones de personas entre hombres y mujeres mueren aproximadamente por accidentes y enfermedades ocupacionales, (6,7)manifiesta que unos de los mayores problemas son la perdida de personas a consecuencia de los accidentes y muchos de ellos llegan a tener enfermedades ocupacionales, (8) afirma que en el mundo muchos trabajadores fallecen como consecuencia de los accidentes en el trabajo y quedan discapacitados por lesiones ocupacionales. (9,10) investigo que en China existe muchos accidentes de trabajo causando 16 795 muertos y 9183 heridos de un reporte total de 877 accidentes y además (11,12) menciona que desde el 2010 al 2019 murieron 7275 muertes en la industria de construcción.

El sector construcción de Malasia no es diferente; los registros del Departamento de Seguridad y Salud Ocupacional (DOSH) indican que el sector de la construcción tiene la mayor cantidad de muertes con aproximadamente la mitad de todos los accidentes ocupacionales (13). En 2021, el DOSH notificó 151 muertes, donde el 37 % de los incidentes fueron incidentes de construcción y el 28 % en el sector manufacturero. Las caídas fatales desde alturas contribuyen en gran medida a ese número según un estudio de Zer-mane et al. (14,15).

En Perú el Ministerio del trabajo reporto 3547 accidentes de este total el 98,33% pertenece a accidentes no mortales y un 0,61% pertenece a accidentes mortales, el 1,03% a incidentes peligrosos y el 0,03% a enfermedades ocupacionales, la consecuencia del accidente se da por accidentes con lesiones, accidentes con invalidez y accidentes de muerte con 2854 varones y 693 mujeres (16,17). Las enfermedades ocupacionales y accidentes de trabajo están directamente relacionadas al entorno laboral y actividades de alto riesgo, los sectores económicos con alta frecuencia de accidentes son los relacionados a manufactura, agricultura, construcción, energía e industrial (18,19). Aprender del pasado siempre ha desempeñado un papel clave en el estímulo a la innovación y a la promoción de avances. Sin duda, las equivocaciones pertenecen a nuestra naturaleza humana, pero todos tenemos capacidades importantes para aprender de ellos. Sin embargo, derivar la recuperación y aplicar los conocimientos adquiridos para evitar errores recurrentes no es tan trivial como puede parecer.

La historia tiende a repetirse, y las lecciones pueden ser ignoradas u olvidadas (20,21). Las diferentes actividades humanas tienen una tolerancia diferente a los errores. Dentro de la industria química, se han realizado esfuerzos significativos para evitar errores y garantizar operaciones seguras. Sin embargo, antes de la segunda mitad de los años sesenta, las palabras "seguridad de los procesos" y "prevención de pérdidas" no existían (22); la manipulación y el almacenamiento de sustancias peligrosas estaban regulados por la seguridad laboral tradicional y las buenas prácticas de ingeniería (23).

Dado que la enseñanza de accidentes graves se mira profundamente afectado por factores humanos, se puede argumentar que un alumno artificial sería un buen apoyo para mejorar las oportunidades de aprendizaje. Los algoritmos de aprendizaje automático podrían ser entrenados para vincular las características del accidente (por ejemplo, sustancias y equipos implicados, magnitud de la liberación, densidad de población) con las consecuencias del accidente, por ejemplo, el número de personas involucradas. Estos modelos predictivos serían un medio rápido, eficaz y económico para apoyar decisiones basadas en el riesgo y la seguridad de los procesos. Sin embargo, el estudio de los datos de los accidentes del proceso a través de algoritmos sigue siendo un tema en gran medida inexplorado (24,25).

En este contexto, esta investigación pretende contribuir a este ámbito de la investigación explorando el empleo de los métodos de aprendizaje automático para analizar y extraer conocimientos de los datos históricos de accidentes. Este estudio responde a las necesidades específicas y apremiantes de herramientas para extraer el conocimiento de accidentes pasados, retener y recordar fácilmente dichos conocimientos para su uso futuro.

Metodología

Ámbito de estudio

El estudio fue realizado a cabo con el uso del modelo de Machine Learning para la predicción del accidente de trabajo por actividad económica, con el propósito de tomar acciones preventivas ante los posibles riesgos encontrados o que sean repetitivos. Para ello se fueron tomados datos recopilados de los accidentes de trabajo por actividad económica en los 24 departamentos del Perú, el cual nos sirvió como datos históricos para realizar la investigación (12).(Ver el tipo de actividad en el Anexo 01)

Metodología de la Investigación

Esta investigación fue realizada mediante el pronóstico de accidentes con datos empíricos para realizar el análisis de estos, y mediante el modelamiento con machine learning se pueda realizar una predicción en un determinado tiempo a futuro (26).

Los datos de accidente de trabajo se han considerado dentro del periodo 2016 al 2021 del sector industria manufacturera, el método del pronóstico consiste en establecer rangos de valor proyectado donde se implementa logaritmos, calculados de modelos, para optimizar los riesgos de los accidentes de trabajo ocurridos en todo el país.

Para esta investigación el presente autor debió seguir el procedimiento utilizado por Liu et al. (2016) donde establece 5 procedimientos a seguir: Base de datos, análisis preliminar, hipótesis, verificación y aplicación (20).

a) Base de datos

Los datos fueron extraídos del ministerio de trabajo del gobierno peruano mediante su página web, donde muestran los reportes de los accidentes de trabajo ocurridos en todos los departamentos del Perú a acceso abierto. Es preciso mencionar que el total de accidentes de trabajo es considerado a toda persona que ha tenido lesiones, invalidez o muerte (27).

b) Análisis preliminar

Fue realizada una limpieza y transformación de datos. Una vez seleccionada la variable de estudio se realizó la limpieza de los datos para obtener datos fiables tras un apoyo visual reorganizando las filas y columnas mediante la función de arrastrar y soltar y seleccionar algunos elementos de datos que faltan y de errores lógicos en los datos quedando listo para su análisis estadístico utilizando el programa R studio.(28).

c) Visualización preliminar

Se utilizó gráficos de líneas y barras con tablas para visualizar y comprender mejor el comportamiento del modelo de accidentes de trabajo por actividad económica mediante la técnica de machine learning.

d) Hipótesis

Según (3) fue realizada la prueba de hipótesis para conocer la predicción de los accidentes de trabajo con respecto a las actividades económicas mediante un modelamiento de técnica de machine learning.

e) Verificación Análisis estadístico

Fue analizado estadísticamente y a través de gráficos de dispersión y diagramas de barras y líneas para explicar el comportamiento del modelamiento de accidentes de trabajo por actividad económica mediante la técnica de machine learning (22,29).

Resultados

Datos evaluados en 90 categorías durante 6 años. Fue llevada a cabo una evaluación para precisar si se tiene diferencia significativa entre la cantidad de accidentes anuales del sector industrial en todos los tipos de actividades, pero como un paso previo es de importancia haber realizado un análisis Chi2 para la primera actividad de la tabla durante los 6 años (29,30) .

	<i>X2016</i>	<i>X2017</i>	<i>X2018</i>	<i>X2019</i>	<i>X2020</i>	<i>X2021</i>
1	7	15	8	66	24	48

Después de aplicar el Chi² se obtuvo como resultado Chi-squared test for given probabilities.

Data:

X-squared = 102.5, df = 5, p-value < 2.2e-16

Las hipótesis a comparar son:

H0: no se tiene diferencia entre los Accidentes de las personas en los años 2016-2021.

Ha: Se tiene diferencia entre los Accidentes de las personas en los años 2016-2021.

El p-valor $< 2.2e-16$ es menor al p-valor crítico de 0.05, por lo que se rechaza H0, y concluimos que se tiene diferencia significativa entre los Accidentes de las personas trabajadas en los años 2016-2021 al 95% de confiabilidad.

Al inspeccionar los valores esperados

[1] 28 28 28 28 28 28

Juntamente con los valores observados

[1] 7 15 8 66 24 48

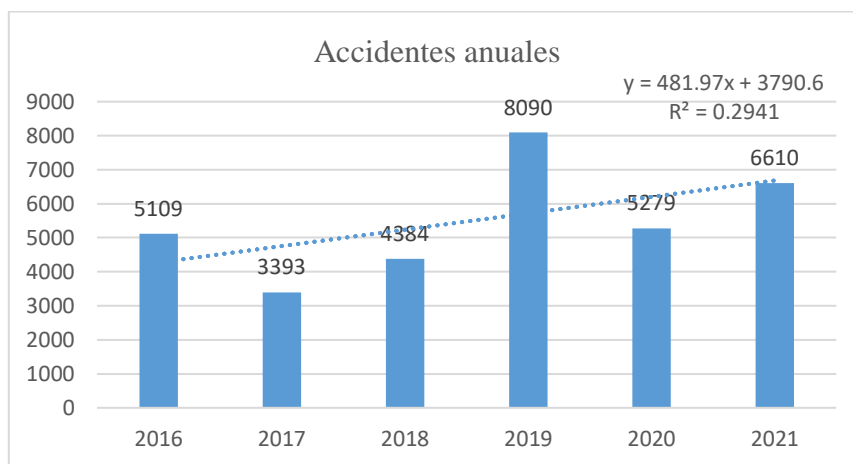


Gráfico N°1: Accidentes anuales.

Fuente: Elaboración propia

Al visualizar los datos de los accidentes anuales podemos proyectar los valores con un modelo de regresión lineal la cual sería

$$Y = 481.97x + 3790.6$$

Con ello se pudo predecir los valores anuales futuros de los accidentes que posiblemente podremos tener.

Por otro lado, fue evidenciada una tendencia al incremento debido a que el coeficiente de la variable accidente (x) es positivo lo que quiere decir que cada año habrá un incremento de 481.97 accidentes.

El número de clusters que este método wss sugiere es de entre 3 a 4.

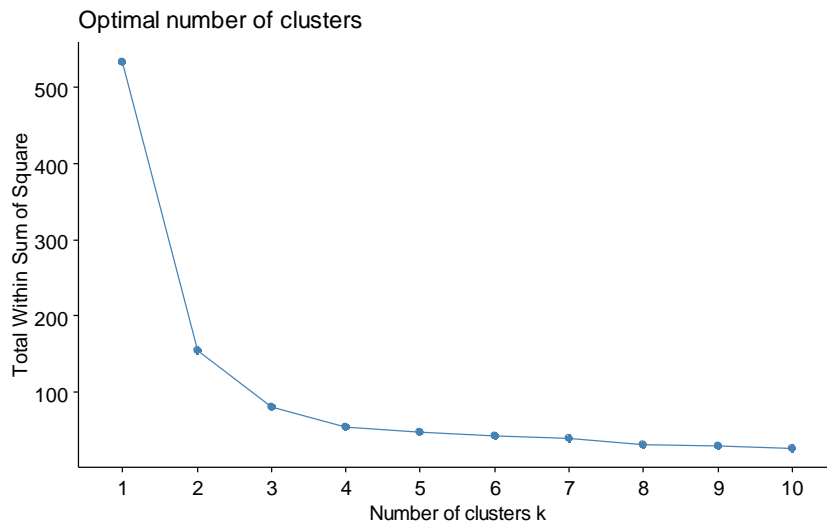


Gráfico N°2: Método WSS

Fuente: Elaboración propia

El número de clusters que este metodo silhouette sugiere es de entre 3 a 4.

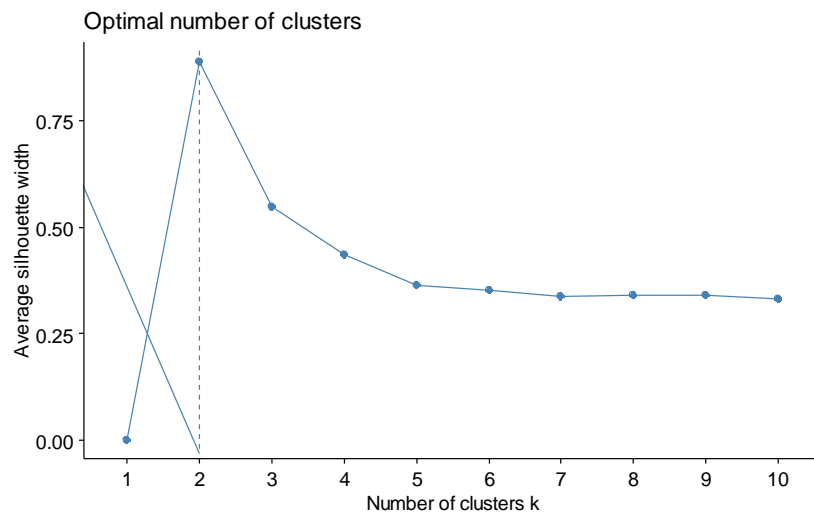


Gráfico N°3: Método silhouette.

Fuente: Elaboración propia

El número de clusters que este metodo gap_stat sugiere es de entre 3 a 4.

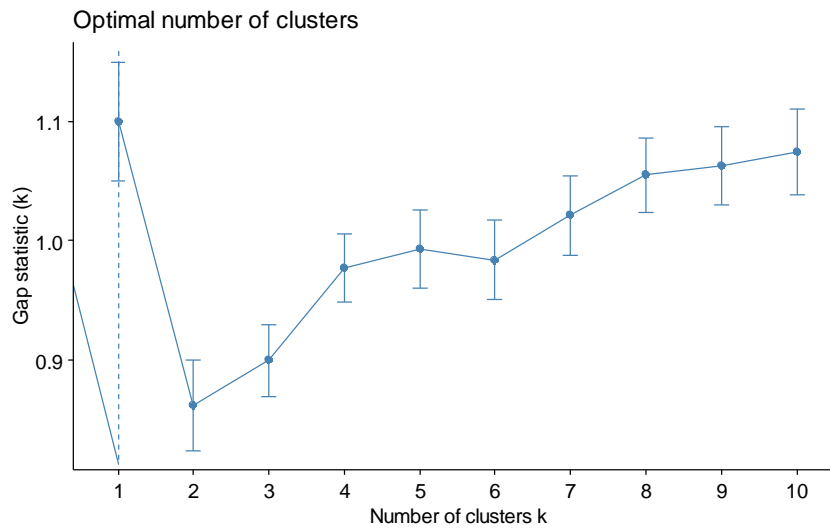


Gráfico N°4: Método Gap_stat.

Fuente: Elaboración propia.

- * Among all indexes:
- * 10 suggested 2 as the best number of clusters
- * 6 suggested 3 as the best number of clusters
- * 4 suggested 4 as the best number of clusters
- * 3 suggested 5 as the best number of clusters
- * 2 suggested 7 as the best number of clusters
- * 3 suggested 10 as the best number of clusters

Según la regla de la mayoría, el mejor número de conglomerados es 2.

Sin embargo, al ser realizado un recuento de las distintas metodologías, sugiere que podamos tener 2 clústeres.

Al llamar a los clusters obtenemos los índices mayor y menor de los mismos.

Cluster means:

	X2016	X2017	X2018	X2019	X2020	X2021
1	-0.1281161	-0.1279747	-0.1285666	-0.1276151	-0.1231747	-0.1228987
2	5.6371079	5.6308864	5.6569310	5.6150631	5.4196887	5.4075432

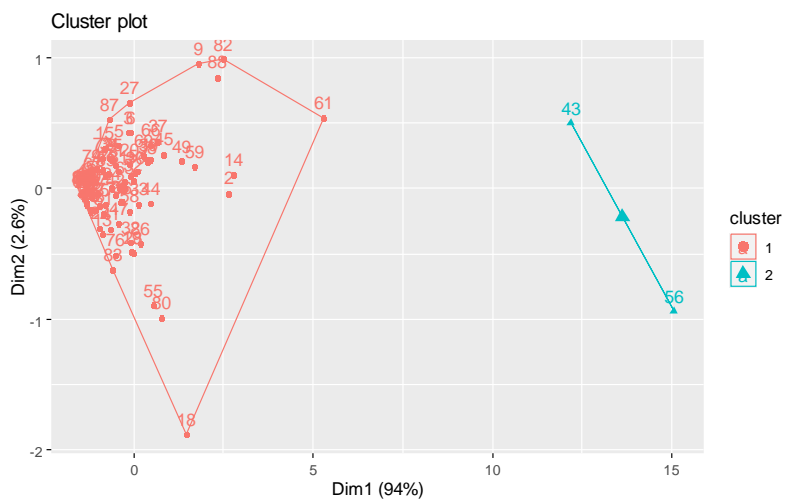


Gráfico N°5: Cluster Euclidean.

Fuente: Elaboración propia

En la presente figura de cluster Euclidean, muestra el comportamiento de los dos clusters totalmente diferentes los cuales indican que deben ser los accidentes que más ocurrencia tienen y los que no suelen ocurrir con frecuencia.

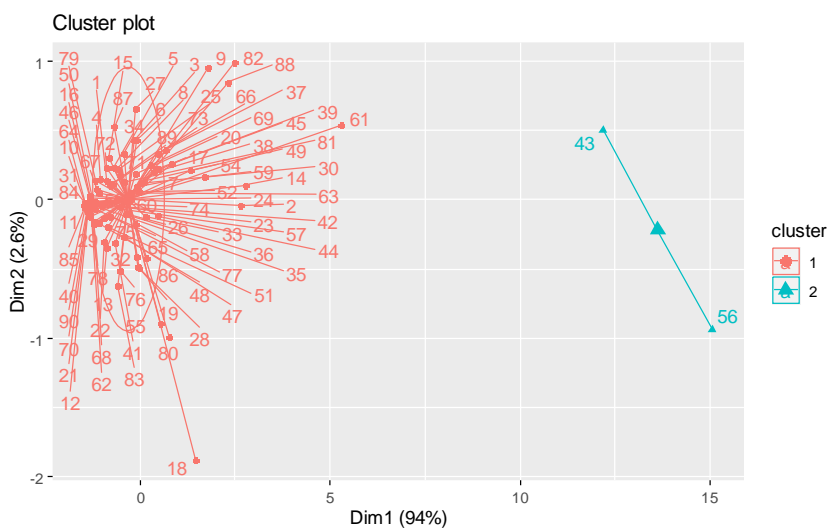


Gráfico N°6: Cluster con distanciamiento elíptico.

Fuente: Elaboración propia

Luego es posible visualizar los mismos clusters pero con distanciamiento elíptico para poder visualizarlo de manera individual como por ejemplo que la actividad 18 se aleja más del centro ortogonal. Y que esa actividad puede ser objeto de estudio con mayor detalle.



Gráfico N°7: Cluster con buena distribución.

Fuente: Elaboración propia

En el siguiente gráfico los conglomerados están más claros y las actividades que tienen menor número de accidentes son las que se acercan más al centroide, es como decir que el centro rojo es como el que menor accidente de trabajo tiene, y se puede afirmar que pueden ser que son las actividades con menor índice de riesgos.

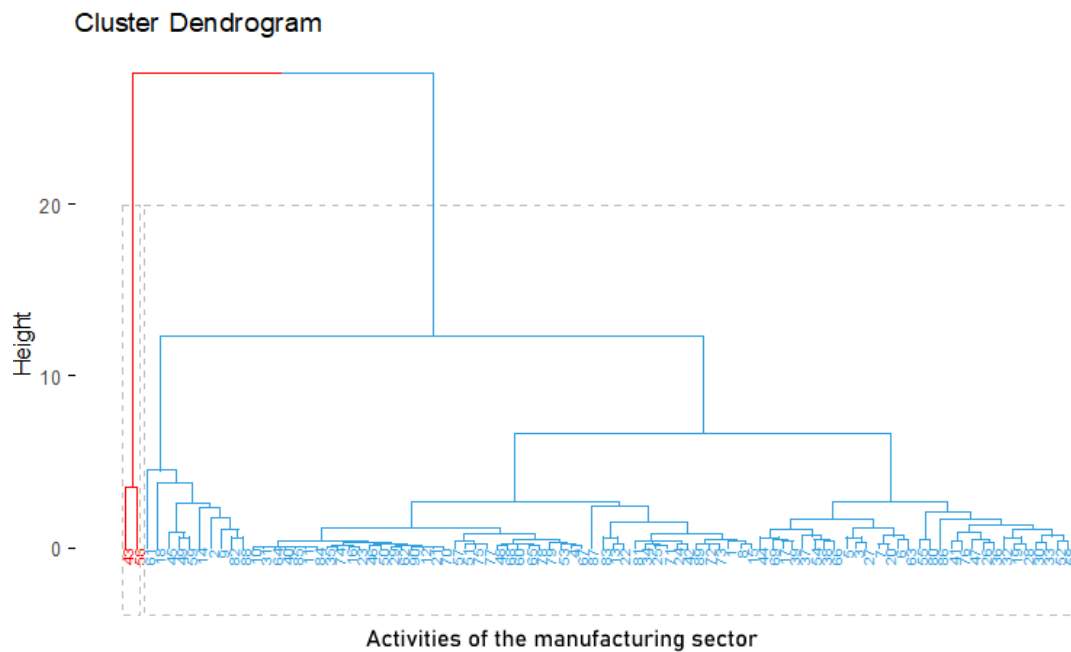


Figura N°1: Clustering tipo jerárquico.

Fuente: Elaboración propia

Esta figura con clustering de tipo jerárquico tiene mayor explicación para la conglomeración de actividades con mayor frecuencia de accidentes. Con ello se puede afirmar que los cluster de color azul representan las actividades con certeza absoluta de accidentes mayoritarios y los cluster restantes las actividades con una incidencia menor o nula de accidentes que podrían afectar a los trabajadores involucrados.

Cluster

	X2016	X2017	X2018	X2019	X2020	X2021
<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	44.8	29.3	38.3	70.8	47.7	59.6
2	582.	406.	507	928	542.	684.

La tabla indica la media de los conglomerados de manera real, es decir que son las medias de las actividades con más ocurrencias y las que menos ocurrencia.

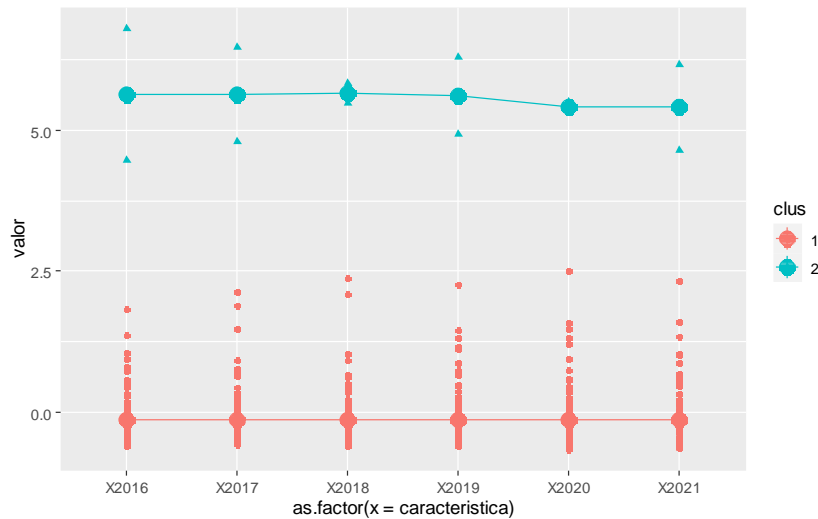


Gráfico N°8: Clusters.

Fuente: Elaboración propia

Finalmente, basado en los dos clusters se puede afirmar que durante los años del 2016 al 2021 se han sucedido los accidentes de trabajo de acuerdo a las diferentes actividades económicas las cuales se reflejan en el Gráfico N°8, lo que indica que el clúster número 1 son las actividades con menor porcentaje de accidentes y los clústeres número 2 son las actividades con mayor número de accidentados.

Resultados

De acuerdo al estudio realizado de accidentes de trabajo en el sector manufacturero con datos trabajados de los años del 2016 al 2021 se obtuvo una ecuación de regresión lineal $Y = 481.97x + 3790$ con lo que se predijo que cada año habrá un incremento de 481.97 accidentes (30,31).

Así también de acuerdo a los gráficos de clústeres se reflejó que se obtuvo mayor porcentaje de accidentes en las actividades de fabricación de productos de plástico y de productos metálicos para la utilización estructural, por lo que se sugiere seguir con la investigación para saber cuáles fueron los factores que se deben a este incremento de accidentes(32).

Por ello se predice que años posteriores habrá incremento de accidentes en el sector manufacturero; así que según los resultados alcanzados se recomienda que deben prestar mayor atención los ingenieros de seguridad en la supervisión debido a que estos accidentes se presentaran con mayor frecuencia y también implementar un sistema de gestión de SST para reducir los accidentes de trabajo en el sector manufacturero (33).

Así también se recomienda que para estudiar este tema a mayor detalle sería necesaria mayor información en cuanto a la cantidad específica de cada tipo de accidente por actividades en la industria manufacturera (34).

Discusión

De acuerdo con el resultado mediante la ecuación de regresión lineal el coeficiente de la variable accidente (x) que se obtuvo salió positivo por lo que se predice que cada año incrementarían los accidentes. Caso contrario a lo que fue obtenido en la empresa PT Semen Tonasa, se obtuvo una tendencia decreciente cada año ya sea en los turnos de trabajo, edad de los trabajadores y tipo de accidentes esto se puede deber a distintas razones ya sea por una buena gestión de seguridad o porque no se han declarado o reportado los accidentes en las áreas.

Así también, este estudio revela que los accidentes laborales y las lesiones en el sector construcción de la ciudad de Yedda son causados principalmente por la ausencia de concienciación y experiencia de los trabajadores, la falta de máquinas y equipos adecuados, la falta de equipos de protección individual convenientes, la carencia de formación y la ausencia de supervisores de seguridad cualificados. Estos factores contribuyen a entornos de trabajo inseguros y resultan en un alto número de lesiones y accidentes. Las caídas de altura, las descargas eléctricas y las insolaciones son los tipos más frecuentes de lesiones y accidentes en este sector. Es fundamental implementar medidas eficaces de control y seguridad en estos lugares de trabajo para reducir o eliminar las causas de lesiones y accidentes (35).

Por otro lado, también se obtuvo un resultado similar a la investigación en cuanto a accidentes laborales en el mismo sector al comparar los años 2018 y 2019. Aunque hubo una variación negativa del -4.6% en las notificaciones de accidentes de trabajo, esta disminución se reflejó precisamente en los meses de diciembre. Sin embargo, al analizar la tendencia de los datos, se identifica una correlación efectiva alta entre el tiempo y la cantidad de accidentes notificados. Esto indica que con el paso del tiempo, se ha visto un incremento significativo en el número de accidentes laborales notificados en el país. Es importante destacar que se menciona que una variación en el tiempo involucra un aumento por arriba del 70% en cantidad de accidentes laborales notificados, lo que sugiere un aumento sustancial en la incidencia de estos eventos a lo largo del tiempo.

En otros países como Ecuador revela varios hallazgos significativos, en primer lugar, se observa una disminución en el número de accidentes laborales calificados particularmente de 2016 a 2017 a pesar del aumento de empresas. Esto sugiere que las normativas legales implementadas en Ecuador, han contribuido a minimizar los accidentes laborales.

En contraste, el sector manufacturero y de construcción muestra una evolución persistente en cantidad de empresas, pero decrecimiento considerable en el

número de accidentes calificados. Esto se debe a que en este sector se tienen organizaciones con riesgos medios y altos, lo que implica la obligatoriedad de implementar Sistemas de Gestión de Seguridad y Salud en el Trabajo.

Referencia Bibliográfica

1. OMS. Mejora de la seguridad y salud en el trabajo en las pequeñas y medianas empresas. WHO. 2011;1–84.
2. Gammarano, Rosina. Quick Guide on sources and uses of statistics on occupational safety and health. 2020 [cited 2023 Feb 27]; Available from: www.ilo.org/publns.
3. Sinclair JK, Selfe J, Taylor PJ, Shore HF, Richards JD. Influence of a knee brace intervention on perceived pain and patellofemoral loading in recreational athletes. *Clinical Biomechanics*. 2016 Aug;37:7–12.
4. Motealleh A, Mohamadi M, Moghadam MB, Nejati N, Arjang N, Ebrahimi N. Effects of Core Neuromuscular Training on Pain, Balance, and Functional Performance in Women With Patellofemoral Pain Syndrome: A Clinical Trial. *J Chiropr Med*. 2019 Mar;18(1):9–18.
5. Ahmadi M, Yalfani A, Gandomi F, Rashid K. The Effect of Twelve-Week Neurofeedback Training on Pain, Proprioception, Strength and Postural Balance in Men with Patellofemoral Pain Syndrome: A Double-Blind Randomized Control Trial. Vol. 7, *JRSR*. 2020.
6. Bonacci J, Hall M, Fox A, Saunders N, Shippides T, Vicenzino B. The influence of cadence and shoes on patellofemoral joint kinetics in runners with patellofemoral pain. *J Sci Med Sport*. 2018 Jun;21(6):574–8.
7. Yalfani A, Ahmadi M, Gandomi F, Rashid K. The Effect of Twelve Week Neurofeedback Training on Perceptual Pain Intensity, Fear of Pain, Pelvic Drop, and Dynamic Knee Valgus Index in Men with Patellofemoral Pain Syndrome: A. 2020 Mar;1–15. Available from: <https://www.researchgate.net/publication/359577696>
8. Yalfani A, Ahmadi M. Patients with Patellofemoral Pain Exhibiting Decrease Vertical Ground Reaction Force Compared to Healthy Individuals during Weight Bearing Tasks: A Systematic Reviews and Meta-Analysis. *Iran J Public Health*. 2023 Feb 6;
9. de Oliveira Silva D, Briani R, Pazzinatto M, Ferrari D, Aragão F, de Azevedo F. Vertical Ground Reaction Forces are Associated with Pain and Self-Reported Functional Status in Recreational Athletes with Patellofemoral Pain. *J Appl Biomech*. 2015 Dec;31(6):409–14.

10. Cruz I, Huerta-Mercado R. Occupational Safety and Health in Peru. *Ann Glob Health*. 2015 Jul 1;81(4):568–75.
11. İşsever H, Ezirmik E, Öztan G, İşsever T. STANDARDIZATION OF WORK ACCIDENTS AND OCCUPATIONAL DISEASES INDICATORS OF SOCIAL SECURITY INSTITUTION BETWEEN 2008-2017 YEARS. *Istanbul Tıp Fakültesi Dergisi*. 2020 Oct 29;83(4).
12. Ministerio de Trabajo y Promoción del Empleo. Documento Elaborado por: Ministerio de Trabajo y Promoción del Empleo. LIMA; 2022 May.
13. Philippos. Ambientes de Trabajo Saludables: un modelo para la acción-OMS [Internet]. OMS. 2010 [cited 2023 Dec 12]. 1–32 p. Available from: https://www3.paho.org/hq/dmdocuments/2010/Healthy%20Workplaces_Print%20Version_SPANISH.pdf
14. Xu Q, Xu K. Analysis of the Characteristics of Fatal Accidents in the Construction Industry in China Based on Statistical Data. *Int J Environ Res Public Health*. 2021 Feb 23;18(4):2162.
15. Ye Y, Zhang S, Rao J, Wang H, Li Y, Wang S, et al. Analysis of National Major Work Safety Accidents in China, 2003-2012 [Internet]. Vol. 45, *Iran J Public Health*. 2016. Available from: <http://ijph.tums.ac.ir>
16. Yoon JH, Hong JS, Roh J, Kim CN, Won JU. Dose - Response relationship between noise exposure and the risk of occupational injury. *Noise Health*. 2015 Jan 1;17(74):43–7.
17. Chong HY, Low TS. Accidents in Malaysian Construction Industry: Statistical Data and Court Cases. *International Journal of Occupational Safety and Ergonomics*. 2014 Jan 8;20(3):503–13.
18. Faquih T, van Smeden M, Luo J, le Cessie S, Kastenmüller G, Krumsiek J, et al. A Workflow for Missing Values Imputation of Untargeted Metabolomics Data. *Metabolites*. 2020 Nov 26;10(12):486.
19. Li H, Fan H, Mao F. A Visualization Approach to Air Pollution Data Exploration—A Case Study of Air Quality Index (PM2.5) in Beijing, China. *Atmosphere (Basel)*. 2016 Feb 29;7(3):35.
20. Liu X, Zhu H, Hu Y, Feng S, Chu Y, Wu Y, et al. Public's Health Risk Awareness on Urban Air Pollution in Chinese Megacities: The Cases of Shanghai, Wuhan and Nanchang. *Int J Environ Res Public Health*. 2016 Aug 25;13(9):845.
21. Szer I, Szer J, Kaszubska M, Miszczak J, Hoła B, Błazik-Borowa E, et al. Influence of the seasons on construction site accidents. *Archives of Civil Engineering*. 2021;67(3):489–504.

22. Trávníček P, Tichá Z, Kotek L. Statistical and text analysis of major accidents to the environment in European Union. *Process Safety Progress*. 2021 Dec 23;40(4):367–74.
23. Tamascelli N, Solini R, Paltrinieri N, Cozzani V. Learning from major accidents: A machine learning approach. *Comput Chem Eng*. 2022 Jun 1;162.
24. Lee JY, Yoon YG, Oh TK, Park S, Ryu S II. A Study on Data Pre-Processing and Accident Prediction Modelling for Occupational Accident Analysis in the Construction Industry. *Applied Sciences*. 2020 Nov 9;10(21):7949.
25. Ahmed S, Hossain MA, Ray SK, Bhuiyan MMI, Sabuj SR. A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance. *Transp Res Interdiscip Perspect*. 2023 May 1;19.
26. Prajapati G, Kumar L, Rekha Patil SS. Road Accident Prediction Using Machine Learning [Internet]. INDIA; 2023 May. Available from: www.jsrtjournal.com
27. Rana V, Joshi H, Parmar D, Jadhav P, Kanojiya M. Road Accident Prediction using Machine Learning Algorithm. *International Research Journal of Engineering and Technology* [Internet]. 2008;3494. Available from: www.irjet.net
28. Christian B, Torrico H, Ariel E, Gonzales R, Erick ;, Gonzales AR. Metodología científica para la realización de investigaciones de mercado e investigaciones sociales cuantitativas Scientific methodology for conducting marketing research and quantitative social research. 2018.
29. Dumont D, Rafael J, Mansilla S, Lorelei S, Martinez S, Nanzy R, et al. Accidentes laborales en el Perú: Análisis de la realidad a partir de datos estadísticos. *Revista Venezolana de Gerencia* [Internet]. 2020;312–29. Available from: <https://orcid.org/0000-0001-9249-3241>.
30. Espinoza-Guano M, Ramos-Guevara J. Análisis comparativo de la accidentabilidad laboral en Ecuador: periodo 2014 al 2019. *593 Digital Publisher CEIT*. 2021 Nov 3;6(6):49–58.
31. Abukhashabah E, Summan A, Balkhyour M. Occupational accidents and injuries in construction industry in Jeddah city. *Saudi J Biol Sci*. 2020 Aug;27(8):1993–8.
32. Zermene A, Mohd Tohir MZ, Zermene H, Baharudin MR, Mohamed Yusoff H. Predicting fatal fall from heights accidents using random forest classification machine learning model. *Saf Sci* [Internet]. 2023 Mar;159:106023. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0925753522003629>

33. Rao Ballamudi K. Road Accident Analysis and Prediction using Machine Learning Algorithmic Approaches. Asian of Journal Humanity, Art and Literature. 2019;6(2).
34. Nai'em MF, Darwis AM, Maksun SS. Trend analysis and projection of work accidents cases based on work shifts, workers age, and accident types. Gac Sanit. 2021;35:S94–7.
35. Rodríguez Giménez M, Ubalde-Lopez M, Cornelio C, López-Ruiz M, G. Benavides F. Evolución de las lesiones por accidentes de trabajo y enfermedades profesionales en la República Argentina, 1997-2015. Arch Prev Riesgos Labor [Internet]. 2019 Jan 1; Available from: http://www.archivosdeprevencion.com/view_document.php?tpd=2&i=12100

Figuras y tablas

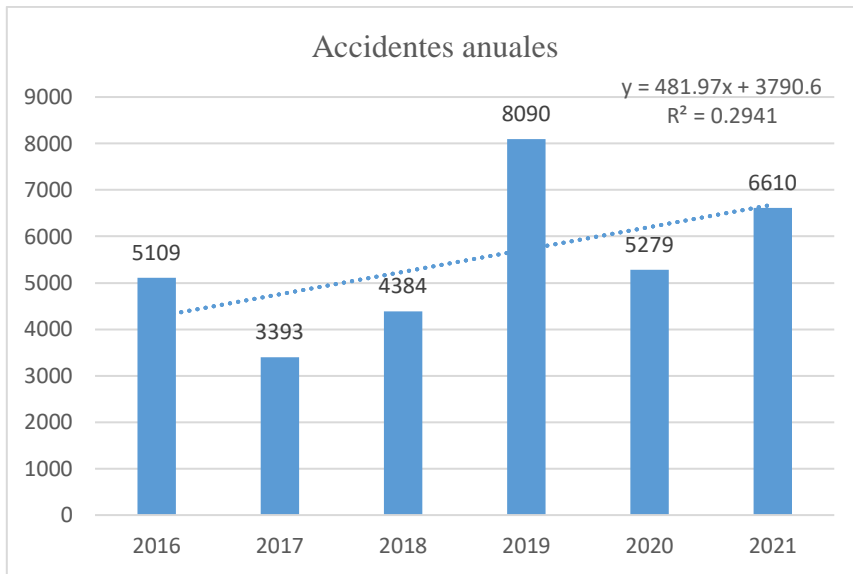


Gráfico N°1: Accidentes anuales.

Fuente: Elaboración propia

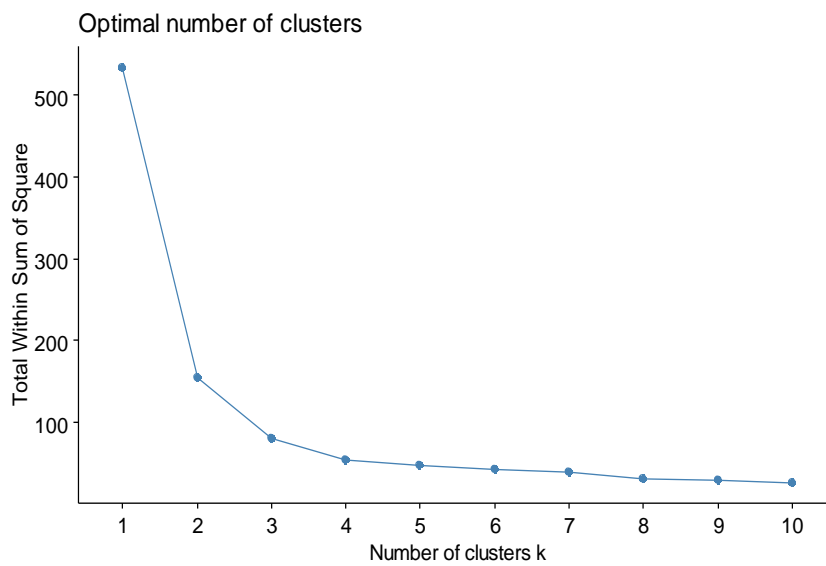


Gráfico N°2: Método WSS

Fuente: Elaboración propia

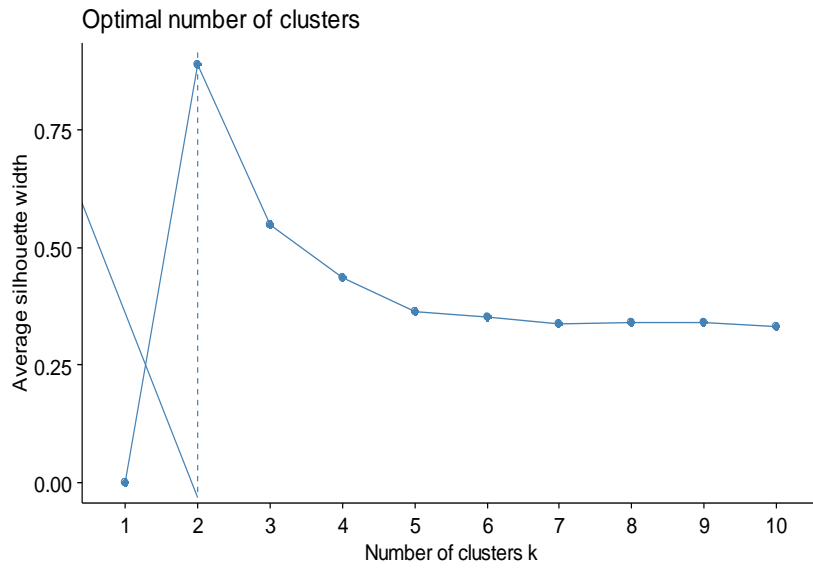


Gráfico N°3: Método silhouette.

Fuente: Elaboración propia

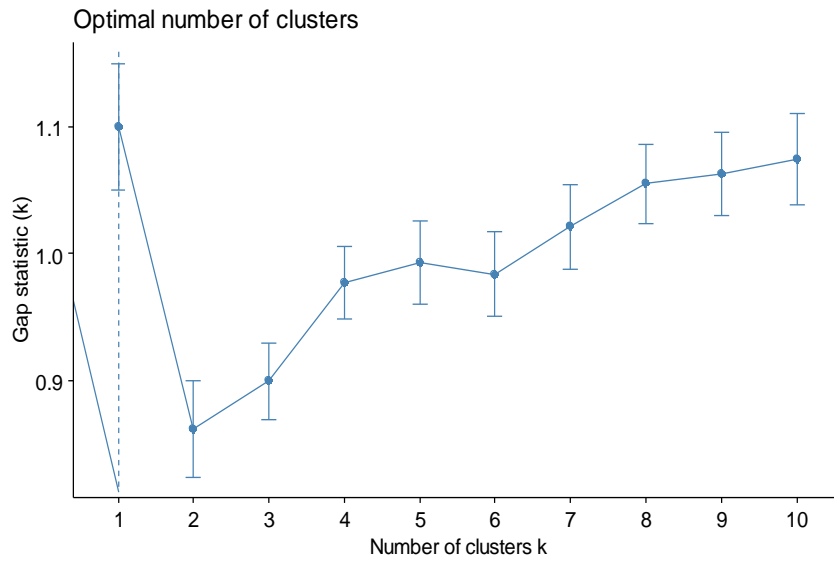


Gráfico N°4: Método Gap_stat.

Fuente: Elaboración propia.



Gráfico N°5: Cluster Euclidean.

Fuente: Elaboración propia

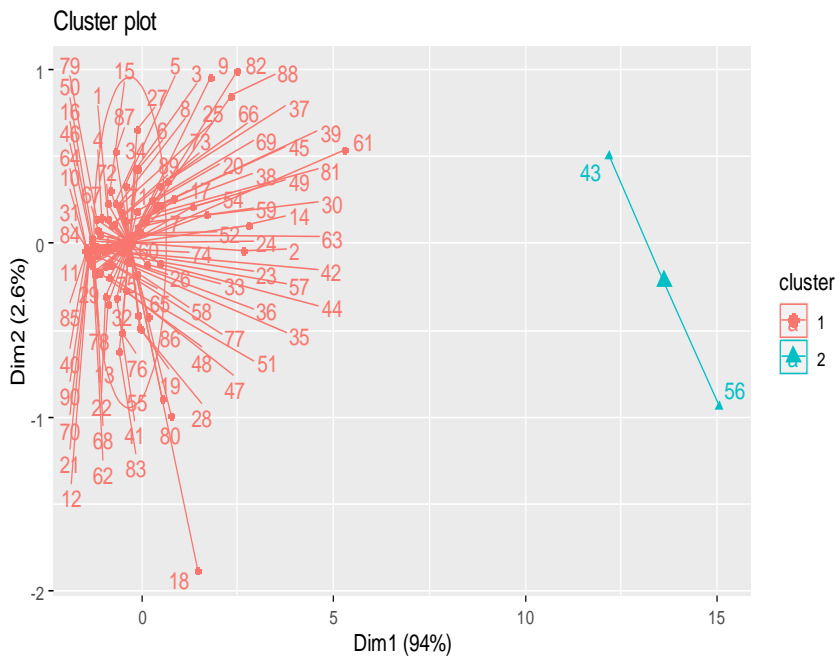


Gráfico N°6: Cluster con distanciamiento elíptico.

Fuente: Elaboración propia

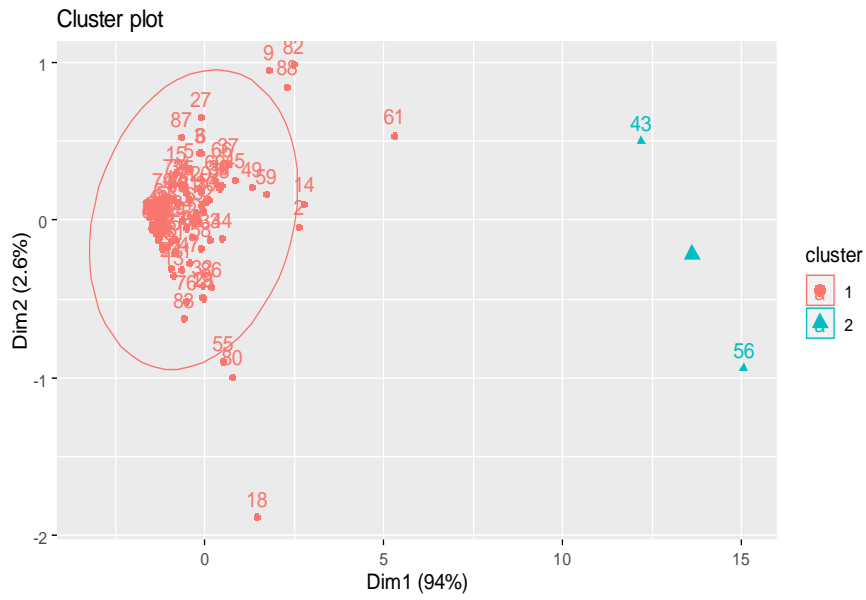


Gráfico N°7: Cluster con buena distribución.

Fuente: Elaboración propia

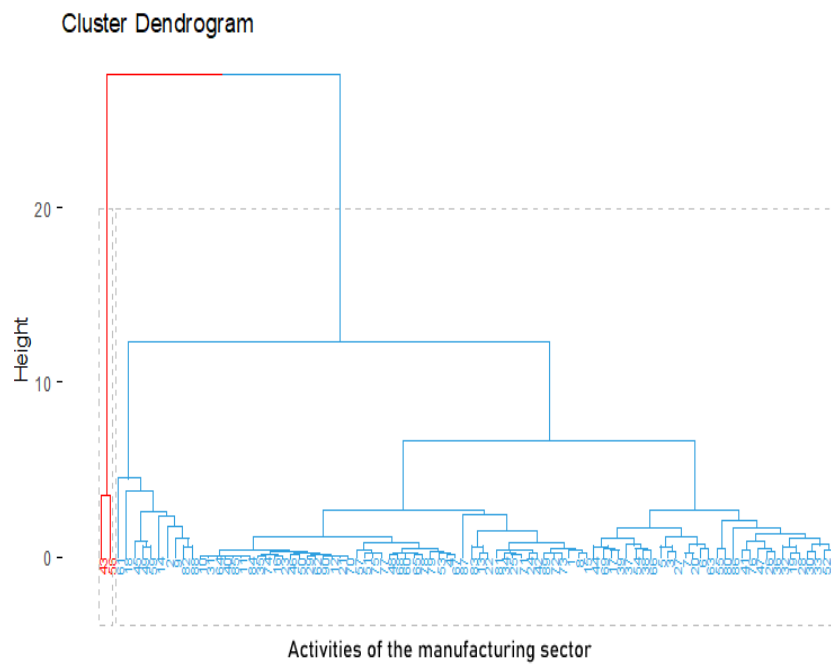


Figura N°1: Clustering tipo jerárquico.

Fuente: Elaboración propia

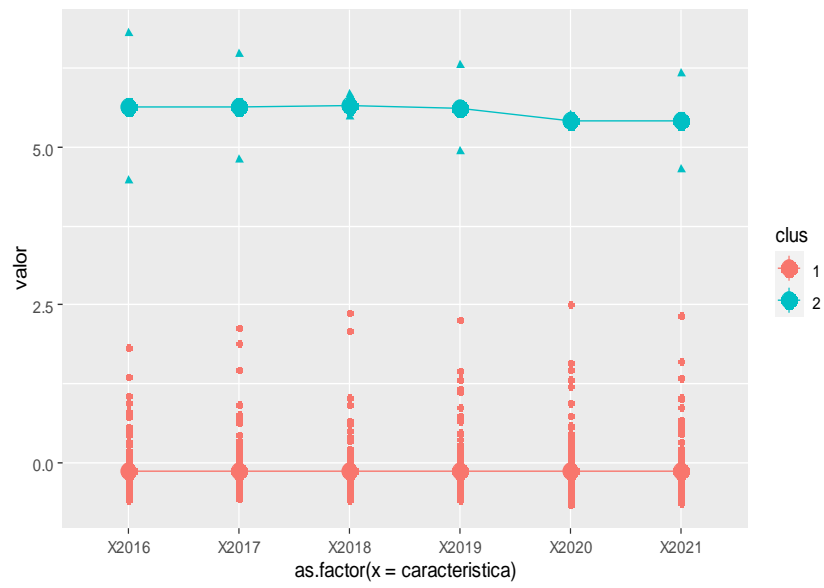


Gráfico N°8: Clusters.

Fuente: Elaboración propia

Anexos

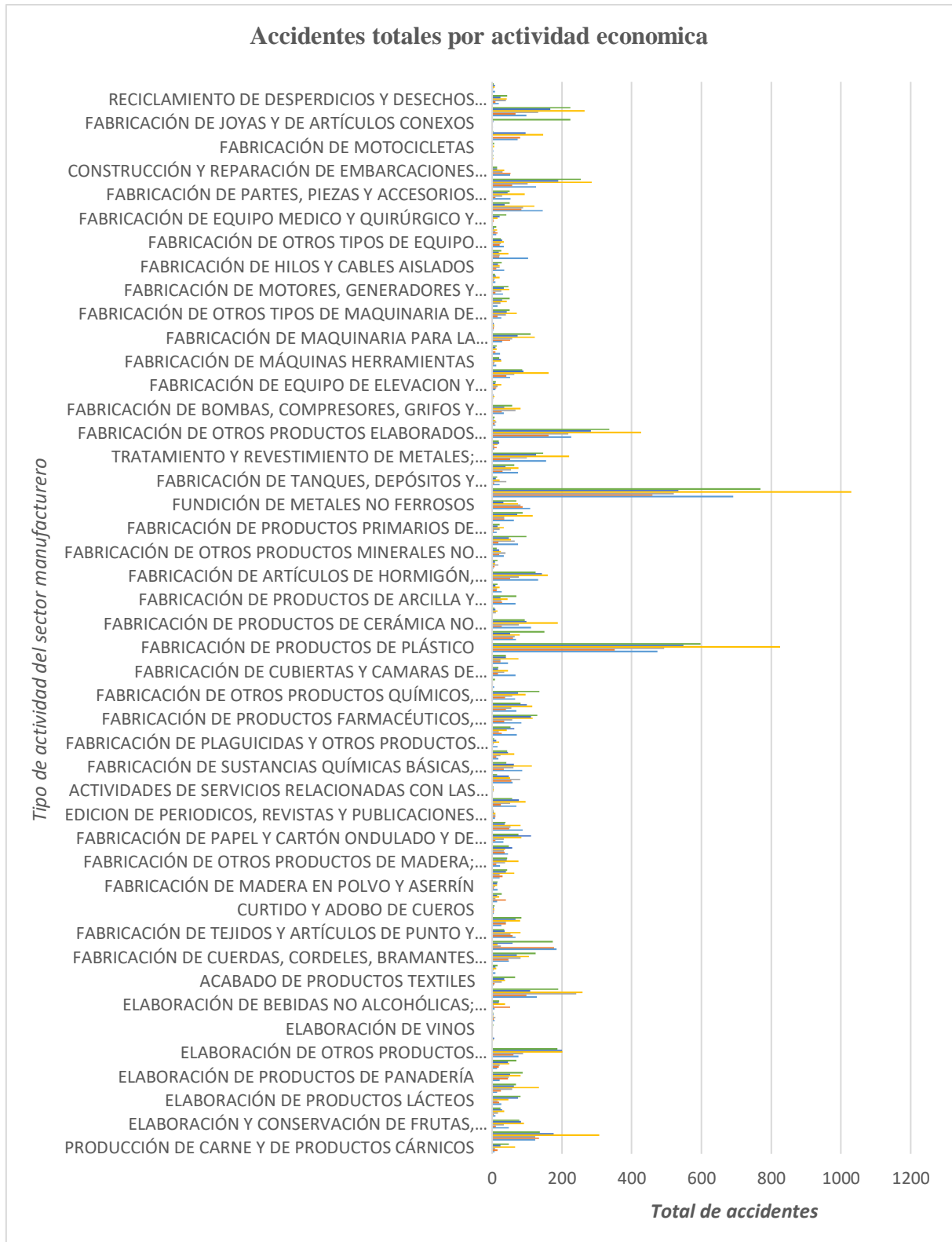


Gráfico 1. Accidentes totales por actividad del sector manufacturero