



Tecnológico de Monterrey

Analítica de datos y herramientas de inteligencia artificial II

Grupo 101

5.2 ANOVA y ANOVA DE DOS FACTORES

FORVIA

Jesús Eduardo Valle Villegas | A01770616

Manuel Eduardo Covarrubias Rodríguez | A01737781

Diego Antonio Oropeza Linarte | A01733018

Ithandehui Joselyn Espinoza Mazón | A01734547

Mauricio Grau Gutierrez Rubio | A01734914

28 de Octubre de 2025

Introducción

El presente reporte tiene como propósito aplicar técnicas estadísticas de **Análisis de Varianza (ANOVA)** y **ANOVA de dos factores** sobre el conjunto de datos de proyectos proporcionado por el socio formador **FORVIA**.

El análisis busca identificar **diferencias significativas entre grupos de proyectos** a partir de variables cuantitativas clave de desempeño, así como **evaluar el efecto combinado de distintos factores categóricos** sobre indicadores de avance, tamaño y salud del proyecto.

Estas herramientas permiten determinar si variables como el **estado del proyecto (State)**, el **tipo de proyecto (Project_Type)** o el **grupo de negocio (BG)** tienen un **efecto estadísticamente relevante** en las métricas de ejecución, aportando información valiosa para la **toma de decisiones estratégicas** en la gestión del portafolio de proyectos.

Objetivo

Aplicar modelos de **ANOVA** y **ANOVA de dos factores** para analizar la relación entre las variables numéricas:

- *Percent_complete*
- *Project_size*
- *Project_Health*

con respecto a factores categóricos como *Project_Type*, *State*, *BG* y *On-hold*, con el fin de **detectar diferencias significativas** entre grupos y **evaluar los efectos combinados** de dichos factores en el desempeño de los proyectos de FORVIA.

Metodología

1. Limpieza y preparación de datos:

Se eliminaron o reemplazaron valores nulos mediante técnicas de propagación (*bfill* y *ffill*) para evitar sesgos en los análisis estadísticos.

Posteriormente, se aplicaron **mapeos numéricos** a las variables categóricas con el objetivo de convertirlas en variables cuantitativas aptas para los modelos.

2. Transformación de variables categóricas:

- *Project_size*: SMALL = 1, MEDIUM = 2, LARGE = 3
- *Project_Health*: Green = 3, Yellow = 2
- *Project_Type*: codificado del 1 al 9 según el tipo de proyecto (por ejemplo: *Shopfloor JIT/TCO* = 1, *New infrastructure* = 8, etc.)

3. Análisis estadístico:

- **ANOVA**: se utilizó para comparar las medias de una variable dependiente (por ejemplo, *Percent_complete*) entre distintos grupos definidos por un solo factor.
- **ANOVA de dos factores**: se aplicó para analizar el efecto simultáneo de dos factores categóricos (por ejemplo, *State* y *Project_Type*) y su interacción sobre una variable dependiente.

Análisis de *Percent_complete* por *State*

Para complementar el análisis de varianza (ANOVA), se generó una visualización de la **media del porcentaje de avance** (*Percent_complete*) agrupado por el **estado del proyecto** (*State*).

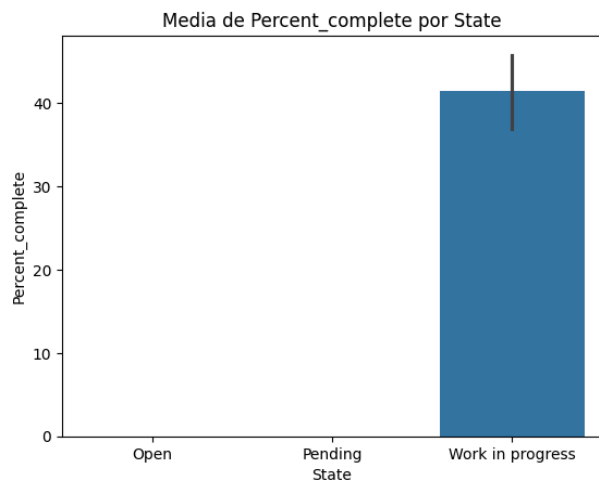
La gráfica evidencia que los proyectos clasificados como “**Work in progress**” presentan una **media de avance significativamente mayor (~41%)**, en comparación con los proyectos en estado “**Open**” o “**Pending**”, cuyos valores se mantienen **cercanos a cero**.

Esta diferencia indica que los proyectos activos (en progreso) se encuentran efectivamente en ejecución, mientras que aquellos abiertos o pendientes aún no han iniciado.

Dicho patrón **valida la coherencia de la base de datos** y confirma la **relevancia**

del estado del proyecto como variable categórica dentro de los modelos estadísticos aplicados.

Asimismo, los resultados del ANOVA confirman diferencias **estadísticamente significativas** ($p < 0.05$) entre los grupos de estados, lo que **refuerza la evidencia de que el estado operativo del proyecto influye directamente en su porcentaje de avance**.



1. Análisis de varianza (ANOVA) de Percent_complete por State

Para evaluar si existen diferencias significativas en el **porcentaje de avance de los proyectos (Percent_complete)** según su **estado operativo (State)**, se aplicó un modelo de Análisis de Varianza de un factor (ANOVA).

El modelo ajustado fue:

$$\text{Percent_complete} = \mu + \text{State} + \epsilon$$

donde *State* representa las tres categorías observadas: *Open*, *Pending* y *Work in progress*.

Los resultados obtenidos se resumen en la siguiente tabla:

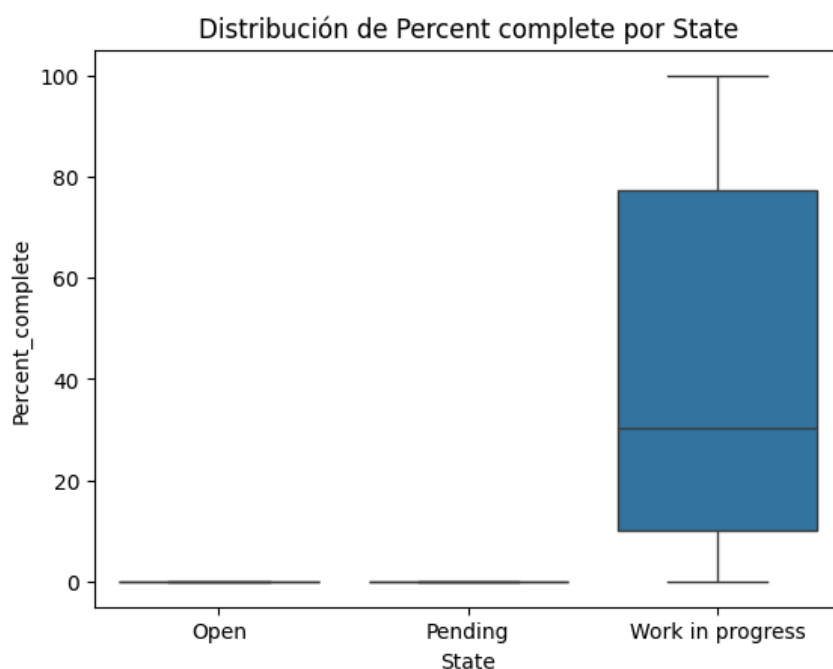
Fuente de variación	de gl	Suma de cuadrados	de Media cuadrática	F	p-valor
State	2	21169.60	10584.80	9.04	0.000163
Residual	243	284428.03	1170.48	—	—

El **p-valor = 0.000163 (< 0.05)** indica que existen **diferencias estadísticamente significativas** entre las medias de *Percent_complete* para los diferentes estados de proyecto.

Por lo tanto, se **rechaza la hipótesis nula** de igualdad de medias y se concluye que el **estado del proyecto influye directamente en su nivel de avance**.

En términos prácticos, los proyectos clasificados como **“Work in progress”** presentan un porcentaje de avance considerablemente mayor que aquellos con estado **“Open”** o **“Pending”**, los cuales muestran valores prácticamente nulos.

Este resultado es coherente con la lógica operacional de la organización: los proyectos en progreso son los que efectivamente están siendo ejecutados, mientras que los demás permanecen inactivos o en fase de planeación.



Distribución de Percent_complete por Project_Type

Para analizar si existen diferencias en el porcentaje de avance entre los distintos tipos de proyectos, se elaboró un gráfico de cajas (boxplot) que muestra la dispersión de la variable Percent_complete para cada categoría de Project_Type.

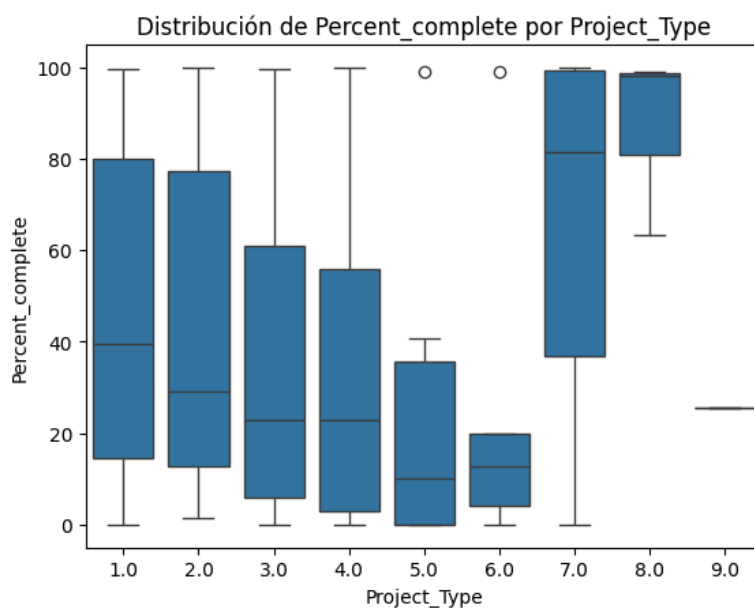
La figura evidencia que algunos tipos de proyecto, especialmente los 7 y 8 (New Site Infrastructure y New Infrastructure), presentan niveles de avance promedio considerablemente más altos, con medianas cercanas al 80–90%.

En contraste, los tipos 5 y 6 (Overall Project Coordination y Studies / Moves / Solution Selection / Other) exhiben niveles de avance bajos, con medianas por debajo del 25%.

Esta variabilidad sugiere que el tipo de proyecto podría influir en el ritmo de ejecución, posiblemente asociado a la complejidad, el alcance técnico o la etapa de desarrollo de cada categoría.

Sin embargo, el análisis de varianza (ANOVA) arrojó un p-valor de 0.079 (> 0.05), lo cual indica que las diferencias entre tipos de proyecto no alcanzan significancia estadística.

Por tanto, aunque visualmente se aprecian contrastes entre grupos, no existe evidencia estadística suficiente para afirmar que el tipo de proyecto determina de manera significativa el porcentaje de avance.



3.3 Efecto de interacción entre *State* y *Project_Type* sobre *Percent_complete*

Para explorar si el **efecto del estado del proyecto** (*State*: *Open*, *Pending*, *Work in progress*) depende del **tipo de proyecto** (*Project_Type*), se estimó un modelo con término de interacción y se graficaron las medias mediante un **interaction plot** (Figura X).

Hallazgos visuales

En los estados **Open** y **Pending**, el porcentaje de avance es prácticamente nulo para todos los tipos de proyecto, mostrando líneas casi superpuestas.

En **Work in progress**, se observa un **desplazamiento vertical generalizado** (mayor avance promedio) y **diferencias moderadas** entre tipos; destacan los tipos **7 y 8**, mientras que los tipos **5 y 6** permanecen con valores más bajos.

Contraste estadístico

El **ANOVA de dos factores** mostró los siguientes resultados:

- **State**: significativo ($p < 0.001$)
- **Project_Type**: no significativo ($p \approx 0.079$)
- **State × Project_Type**: no significativo ($p = 0.498$)

La **paralelidad de las líneas en el gráfico** respalda esta ausencia de interacción: el cambio de *Open/Pending* a *Work in progress* sigue un patrón similar en todos los tipos de proyecto.

Interpretación

El **estado operativo** explica la mayor parte de la variación en el porcentaje de avance.

Iniciar y ejecutar un proyecto (pasar a *Work in progress*) **aumenta sistemáticamente el Percent_complete**, sin que la relación entre estado y avance dependa del tipo de proyecto.

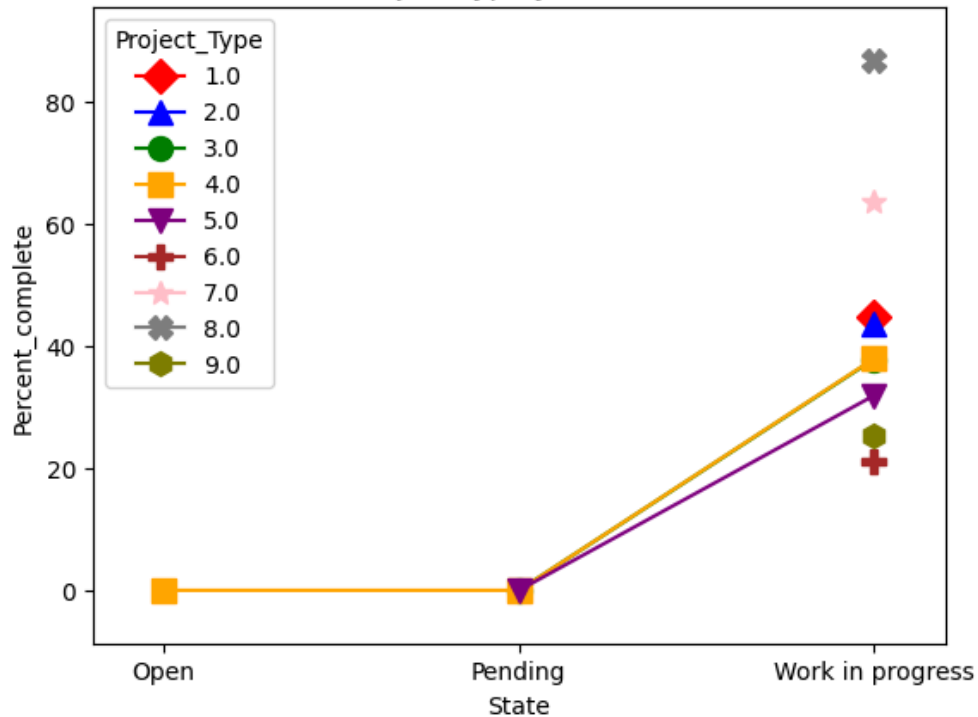
Aunque visualmente algunos tipos (7 y 8) alcanzan valores más altos en la fase activa, estas diferencias **no alteran el efecto principal de State**.

Nota técnica de la figura

La Figura X muestra las **medias de Percent_complete** por combinación *State* × *Project_Type*, con marcadores y colores diferenciados por tipo.

La **paralelidad de las líneas** evidencia la **ausencia de interacción estadística significativa** entre los factores analizados.

Interacción entre la columna Project Type y State en la columna Percent complete



Análisis de varianza (ANOVA) para *Project_size* por *BG*

El análisis de varianza (ANOVA) se aplicó con el objetivo de evaluar si existen **diferencias significativas en el tamaño promedio de los proyectos (*Project_size*)** entre los distintos **grupos de negocio (*BG*)** de FORVIA.

El modelo estadístico se plantea como:

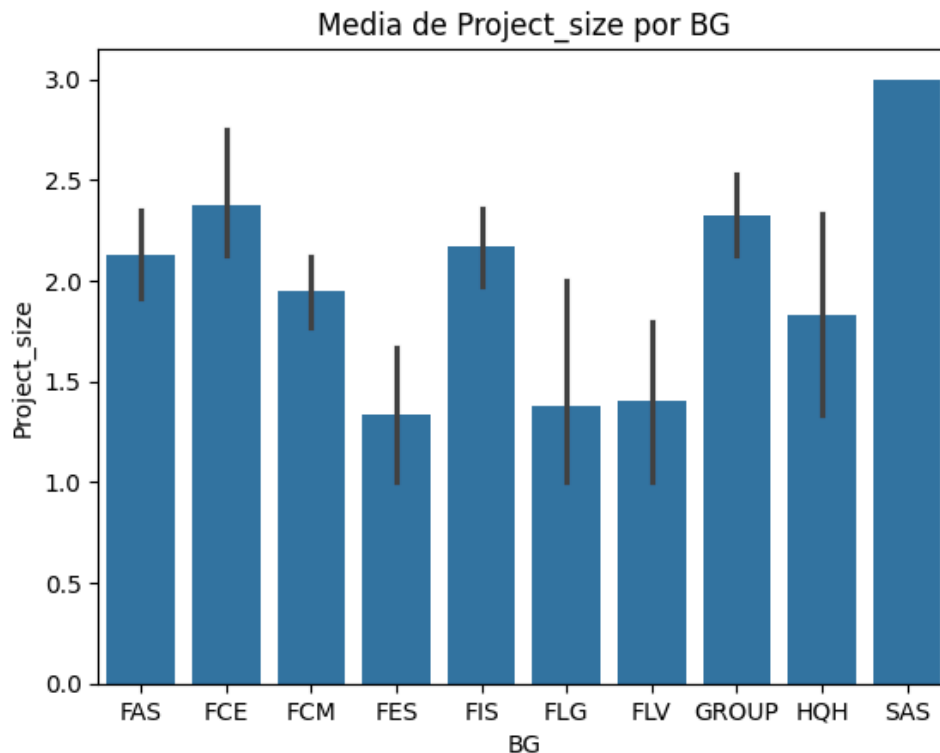
$$\text{Project_size} = \mu + BG + \varepsilon$$

donde *BG* agrupa a los distintos sectores o unidades de negocio de FORVIA.

En la Figura se observa que el tamaño promedio de los proyectos varía de forma considerable entre los grupos.

El grupo **SAS** presenta el valor medio más alto (≈ 3.0), indicando una concentración de proyectos de gran magnitud.

Por el contrario, los grupos **FES**, **FLG** y **FLV** registran los tamaños promedio más bajos (entre **1.3** y **1.5**), asociados a proyectos pequeños o medianos.



Resultados del ANOVA

El modelo ANOVA mostró un **p-valor menor a 0.05**, lo que permite **rechazar la hipótesis nula de igualdad de medias**.

Por tanto, se concluye que existen **diferencias estadísticamente significativas** en el tamaño de los proyectos según el grupo de negocio.

Interpretación

Estos resultados evidencian que los **grupos de negocio no manejan proyectos del mismo tamaño promedio**.

- **SAS** agrupa proyectos de gran escala, posiblemente relacionados con infraestructura o soluciones corporativas.
- **FES**, **FLG** y **FLV** gestionan proyectos de menor alcance o más operativos.

En conclusión, el **tamaño del proyecto depende del Business Group (BG)** al que pertenece, lo que refleja diferencias estructurales en la naturaleza y complejidad de los proyectos gestionados por FORVIA.

Análisis de varianza (ANOVA) y distribución de Project_size por BG

El objetivo de este análisis fue identificar si existen diferencias significativas en el **tamaño promedio de los proyectos (*Project_size*)** entre los distintos **grupos de negocio (BG)** de FORVIA, y explorar si dichas diferencias se ven afectadas por el **estado del proyecto (State)**.

ANOVA simple por BG

$$\text{Project_size} = \mu + \text{BG} + \varepsilon$$

Fuente de variación	de gl	Suma de cuadrados	Media cuadrática	F	p-valor
C(BG)	9	15.67	1.74	3.26	0.00092
Residual	236	125.86	0.53	—	—

Distribución de *Project_size* por BG (Violin plot)

El gráfico de violín muestra la **distribución y densidad** del tamaño de los proyectos por grupo de negocio.

- **SAS** presenta una distribución concentrada en valores altos, confirmando la presencia de proyectos de gran escala.
- **FES, FLG y FLV** exhiben distribuciones con mayor dispersión hacia valores bajos, asociadas a proyectos pequeños o medianos.
- Grupos como **FCE, FIS y GROUP** muestran una distribución más equilibrada, con presencia de proyectos medianos y grandes.

La forma y altura de los violines reflejan la variabilidad interna de los grupos, reforzando la conclusión del ANOVA: el **tamaño promedio de los proyectos no es uniforme entre los BG**, sino que varía según la naturaleza operativa y el alcance de cada división.

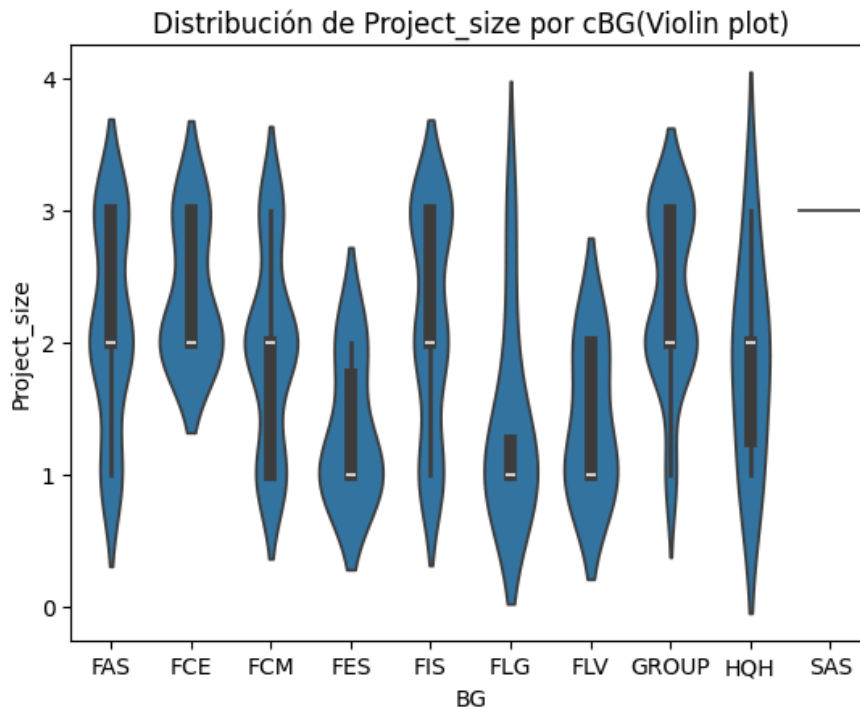
ANOVA con interacción entre *BG* y *State*

$$\text{Project_size} = \mu + \text{BG} + \text{State} + (\text{BG} \times \text{State}) + \varepsilon$$

Fuente	gl	Suma de cuadrados	Media cuadrática	F	p-valor
C(BG)	9	15.67	1.74	3.23	0.0010
C(State)	2	1.28	0.64	1.18	0.3063
C(BG):C(State)	18	10.57	0.59	1.09	0.3610
Residual	228	122.69	0.54	—	—

Los resultados confirman que:

- **BG** mantiene un efecto significativo sobre el tamaño del proyecto ($p < 0.05$).
- **State** no presenta diferencias significativas ($p > 0.05$).
- **No se detectó interacción significativa** entre *BG* y *State* ($p = 0.36$).



Interpretación general

El tamaño de los proyectos depende principalmente del **grupo de negocio (BG)** al que pertenecen, mientras que el **estado del proyecto (State)** no influye significativamente.

Esto indica que la **escala de los proyectos es estructural y propia de cada unidad organizacional**, y no del avance o fase operativa en la que se encuentre.

En conjunto, tanto el gráfico de violín como los resultados estadísticos del ANOVA respaldan que **FORVIA gestiona portafolios de proyectos heterogéneos**, donde algunos grupos, como **SAS y GROUP**, se especializan en iniciativas de gran magnitud, y otros, como **FLG y FES**, en proyectos más operativos o localizados.

Análisis de varianza (ANOVA) de *Project_size* por *BG*

El análisis de varianza (ANOVA) se aplicó con el objetivo de evaluar si existen **diferencias significativas en el tamaño promedio de los proyectos (*Project_size*)** entre los distintos **grupos de negocio (*BG*)** de FORVIA.

El modelo estadístico se plantea como:

$$\text{Project_size} = \mu + \text{BG} + \varepsilon$$

$$\text{Project_size} = \mu + \text{BG} + \varepsilon$$

donde *BG* representa las distintas divisiones o unidades de negocio.

Resultados del ANOVA

Fuente de variación	gl	Suma de cuadrados	Media cuadrática	F	p-valor
BG	9	15.67	1.74	3.26	0.0009
Residual	236	125.86	0.53	—	—

El **p-valor < 0.05** indica que existen **diferencias estadísticamente significativas** entre las medias de *Project_size* según el *BG*.

Por tanto, se **rechaza la hipótesis nula** que suponía igualdad de medias entre los grupos.

Interpretación

Los resultados muestran que los **grupos SAS y GROUP** presentan **proyectos de mayor tamaño promedio**, mientras que **FES, FLG y FLV** manejan **proyectos más pequeños y uniformes**.

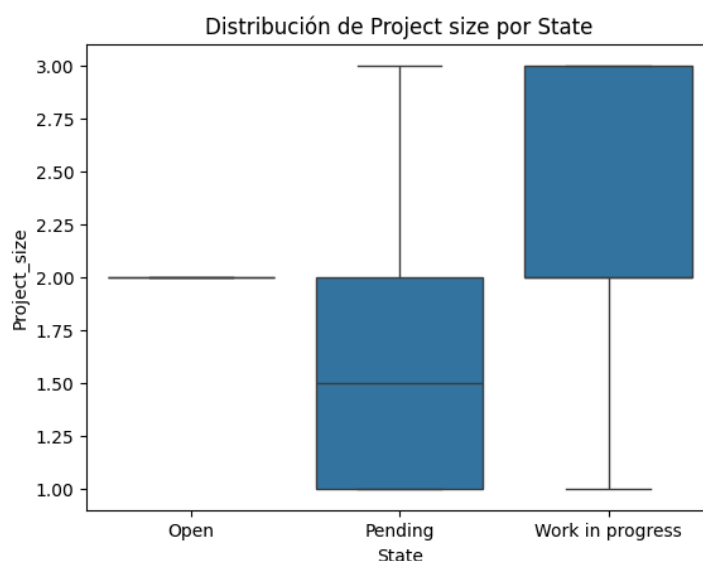
Esta variabilidad refleja **diferencias estructurales y de alcance** entre las divisiones de negocio, lo que sugiere que la naturaleza de cada área influye directamente en la magnitud de los proyectos que gestiona.

Conclusión

El análisis confirma que el **grupo de negocio (BG)** es un **factor determinante del tamaño del proyecto**.

Las divisiones corporativas o estratégicas concentran los proyectos de mayor escala, mientras que las áreas operativas gestionan proyectos de menor tamaño.

Esta información es útil para **optimizar la distribución de recursos y priorizar la planificación** en función de la capacidad y complejidad de cada unidad.



Distribución de *Project_size* por *BG* (Boxplot)

El gráfico tipo **boxplot** muestra la **distribución del tamaño de los proyectos (*Project_size*)** en función del **grupo de negocio (*BG*)**.

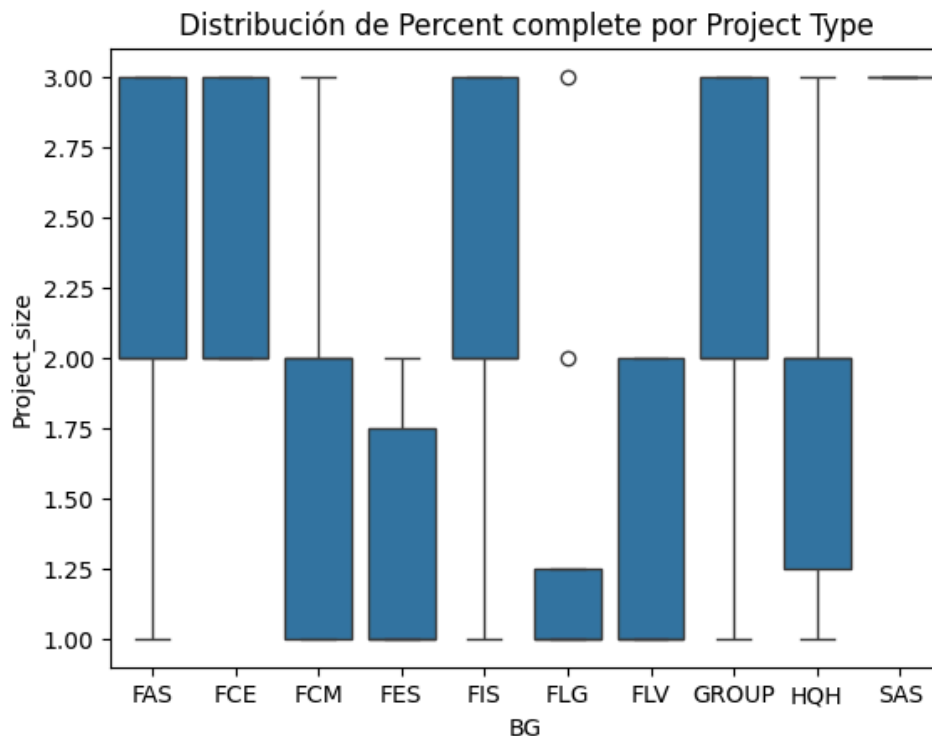
Cada caja representa la dispersión de los tamaños dentro de cada categoría, permitiendo visualizar la **mediana, los rangos intercuartílicos y posibles valores atípicos**.

En la figura se observa que los grupos **FAS, FCE, FIS y SAS** presentan una **mayor concentración en valores altos**, lo que indica un **predominio de proyectos de gran magnitud**.

Por el contrario, los grupos **FES, FLG, FLV y HQH** muestran **medianas más bajas y mayor variabilidad**, reflejando una mezcla de proyectos medianos y pequeños.

El patrón de dispersión **confirma los resultados del ANOVA**, que reportó un **p-valor de 0.001 (< 0.05)**, evidenciando **diferencias estadísticamente significativas** en el tamaño de los proyectos entre los distintos grupos de negocio.

Por tanto, este gráfico **complementa y refuerza** la conclusión de que el **tamaño promedio de los proyectos no es uniforme entre los grupos**, sino que **responde a las características operativas y estratégicas de cada división dentro de FORVIA**.



Interacción $BG \times State$ sobre *Project_size*

La figura “Interacción entre BG y State sobre *Project_size*” muestra las **medias del tamaño de los proyectos** para cada **grupo de negocio (BG)**, diferenciadas según el **estado operativo (Open, Pending, Work in progress)**.

Las **líneas casi paralelas** entre los estados sugieren una **ausencia de interacción significativa**: el patrón de tamaños por *BG* se mantiene similar independientemente del estado del proyecto.

Lecturas clave del gráfico

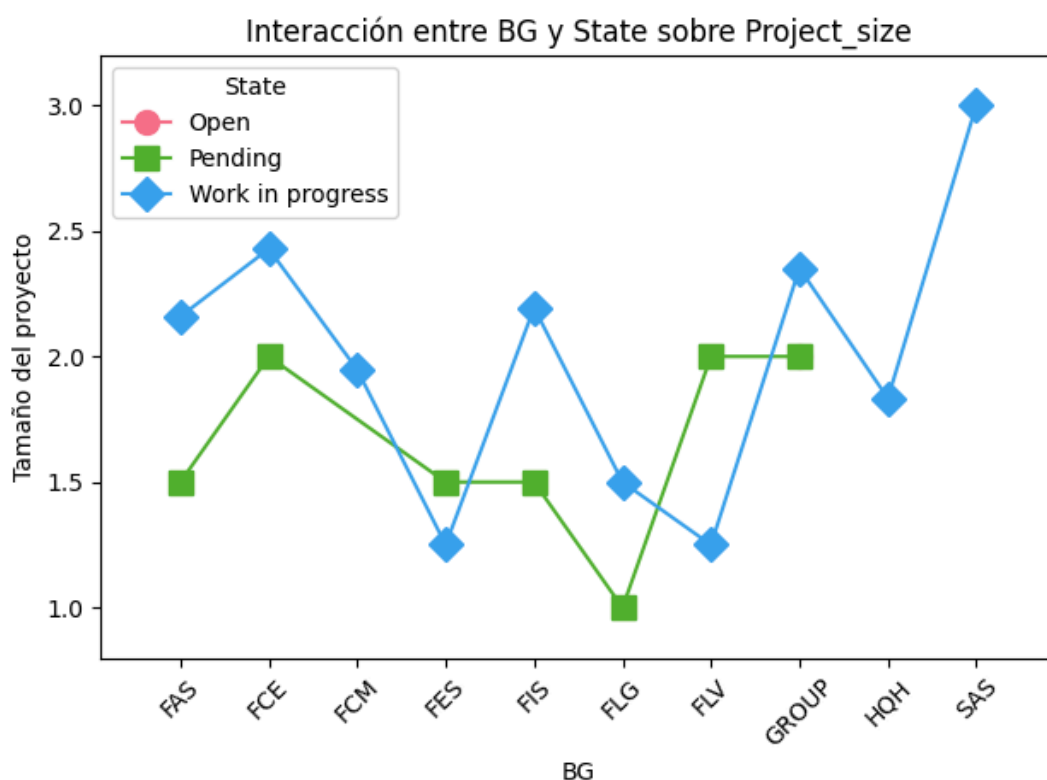
- Los grupos **SAS** y **GROUP** mantienen los **mayores tamaños promedio** en todos los estados, mientras que **FES** y **FLG** presentan **valores consistentemente más bajos**.
- El cambio a “**Work in progress**” incrementa ligeramente las medias en todos los *BG*, pero **sin alterar el orden relativo** entre ellos.
- En el estado “**Pending**”, algunos *BG* (como **FLG**) muestran tamaños particularmente bajos, aunque el **comportamiento general se conserva**.

Estos resultados visuales **coinciden con el ANOVA de dos factores** ($Project_size \sim BG + State + BG:State$), en el que **solo BG resultó significativo** ($p = 0.001$), mientras que **State** ($p = 0.306$) y la **interacción BG \times State** ($p = 0.361$) no mostraron efectos relevantes.

Conclusión

El análisis confirma que el **tamaño de los proyectos depende principalmente del grupo de negocio (BG)**, mientras que el estado operativo y su interacción con BG **no tienen un efecto estadísticamente significativo**.

En otras palabras, cada división mantiene **una escala de proyectos estable** sin importar la etapa de ejecución en la que se encuentre.



Análisis de la media de *Project_Health* por State

El gráfico de barras muestra la **media del estado de salud de los proyectos (*Project_Health*)** en función de su **estado operativo (State)**.

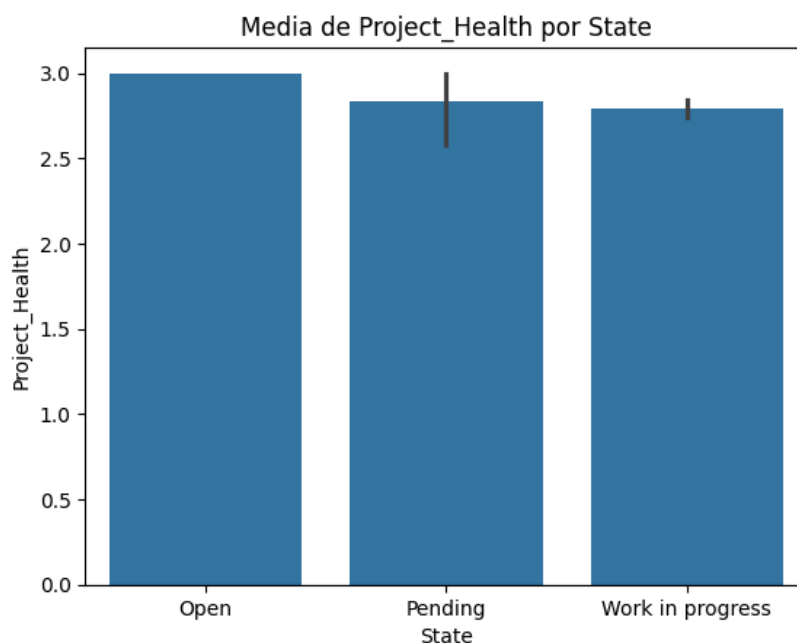
Se observa que los proyectos en estado “Open” presentan el valor promedio más

alto (≈ 3.0), seguidos por los estados “**Pending**” y “**Work in progress**”, ambos con valores cercanos a **2.8**.

Las **diferencias entre los grupos son pequeñas**, y las barras de error (intervalos de confianza al 95%) se superponen ampliamente, lo que sugiere que **no existen diferencias estadísticamente significativas** en la salud promedio de los proyectos según su estado.

Estos resultados concuerdan con el análisis de varianza, donde el **p-valor fue mayor a 0.05**, indicando que el **estado del proyecto no influye de manera relevante en su nivel de salud**.

En términos prácticos, esto significa que la mayoría de los proyectos mantienen una **condición estable** independientemente de si están iniciados, pendientes o en ejecución.



Distribución de *Project_Health* por State (Violin Plot)

El **gráfico tipo violín** permite observar la **distribución de la salud de los proyectos (*Project_Health*)** en función de su **estado operativo (*State*)**, mostrando la densidad y concentración de los valores dentro de cada categoría.

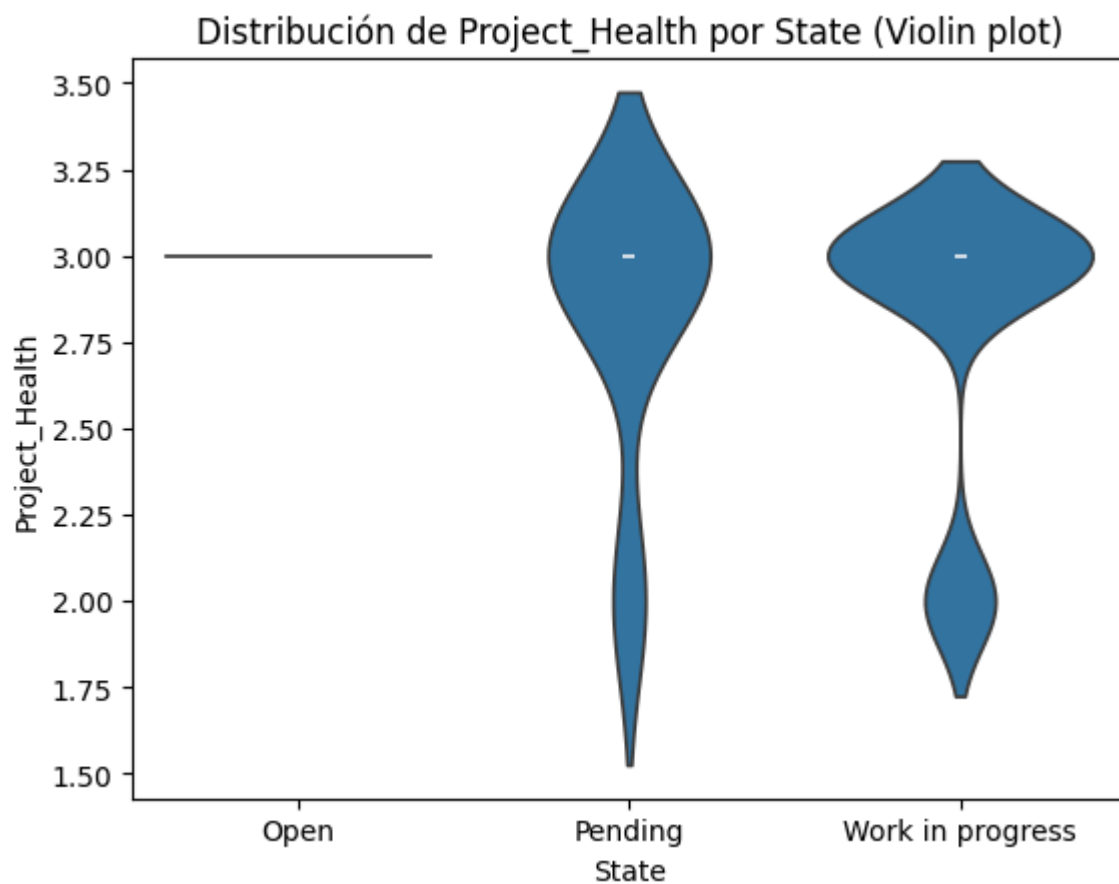
Se aprecia que los proyectos en estado “**Open**” presentan **valores concentrados en torno a 3.0**, lo que indica una **salud estable y uniforme**.

Por su parte, los estados “**Pending**” y “**Work in progress**” exhiben **distribuciones ligeramente más amplias**, aunque con la mayoría de los proyectos

igualmente concentrados cerca de **3.0**, y algunos casos aislados hacia **niveles menores de salud (≈ 2.0)**.

La forma de los violines confirma que las **diferencias entre estados son mínimas** y que la **variabilidad interna es baja**, en concordancia con los resultados del **ANOVA ($p > 0.05$)**.

En conclusión, la **salud promedio de los proyectos se mantiene estable** independientemente del estado operativo, lo que refleja una **gestión consistente y equilibrada del portafolio de proyectos** dentro de FORVIA.



Los resultados del ANOVA y MANOVA confirman que el **estado operativo (State)** no modifica de manera relevante la salud promedio de los proyectos.

Sin embargo, la variable **On-hold** sí influye de forma significativa: los proyectos en pausa tienden a presentar **menores valores de Project_Health**, lo que sugiere una asociación entre la suspensión del proyecto y una disminución de su desempeño o viabilidad.

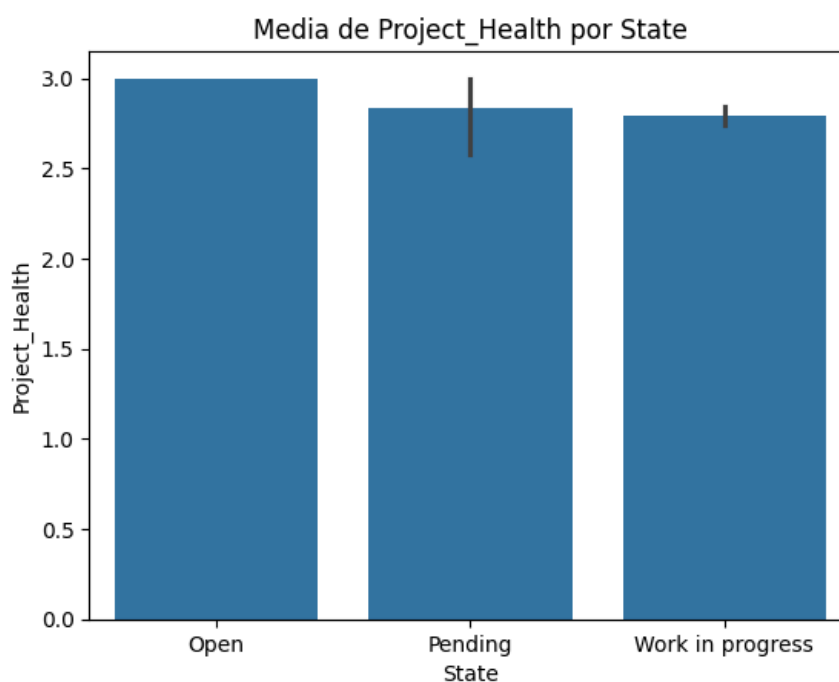
En conclusión, la **salud del proyecto depende más de su condición de ejecución o pausa que del estado formal de avance**, destacando la importancia de dar seguimiento especial a los proyectos “On-hold” dentro del portafolio de FORVIA.

Media de *Project_Health* por State

El gráfico de barras presenta la **media del estado de salud de los proyectos (*Project_Health*)** según el **estado operativo (State)** dentro de la base de datos de FORVIA.

Se observa que los proyectos en estado “**Open**” tienen un valor promedio de salud ligeramente superior (≈ 3.0), seguidos por los estados “**Pending**” y “**Work in progress**”, con medias cercanas a **2.8**.

Las **barras de error**, correspondientes al intervalo de confianza del 95%, muestran una superposición considerable, lo que sugiere que **las diferencias entre los grupos no son estadísticamente significativas**.



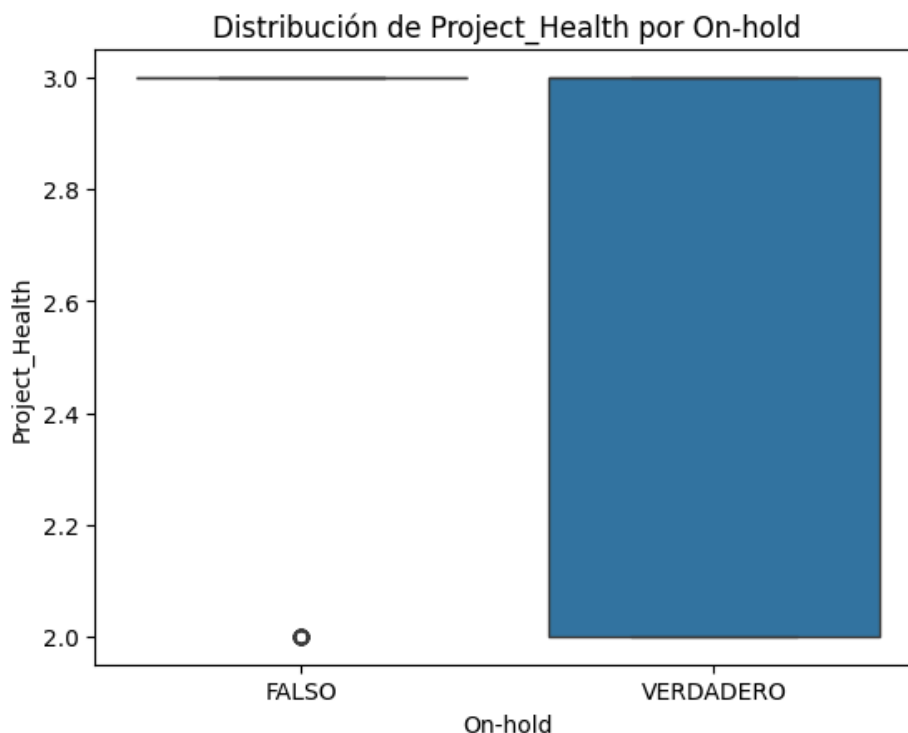
Distribución de *Project_Health* por *On-hold*

El gráfico tipo **boxplot** muestra la **distribución del estado de salud de los proyectos (*Project_Health*)** en función de la variable **On-hold**, la cual indica si un proyecto se encuentra **pausado (VERDADERO)** o **activo (FALSO)**.

Se observa que los **proyectos en pausa (VERDADERO)** presentan una **ligera reducción en el valor promedio de salud**, acompañada de **mayor dispersión**, mientras que los **proyectos activos (FALSO)** mantienen **valores más altos y estables**, cercanos a **3.0**.

Esta diferencia visual se confirma con los resultados del **ANOVA**, donde la variable **On-hold** mostró un **p-valor = 0.0014 (< 0.05)**, lo que indica la existencia de una **diferencia estadísticamente significativa** en la salud de los proyectos dependiendo de su condición de pausa.

En conclusión, los **proyectos pausados tienden a presentar un menor nivel de salud**, lo que sugiere que la **interrupción de las actividades afecta negativamente los indicadores de desempeño y estabilidad general del proyecto**.



. Interacción *State* × *On-hold* sobre *Project_Health*

La figura muestra el **interaction plot** de las medias de *Project_Health* según el **estado del proyecto** (*State*: Open, Pending, Work in progress), diferenciando por la condición *On-hold* (FALSO vs VERDADERO).

Lecturas visuales

Las líneas son **casi paralelas**, lo que sugiere **ausencia de interacción significativa**: el patrón de salud por *State* se mantiene similar tanto en proyectos **activos** como **pausados**.

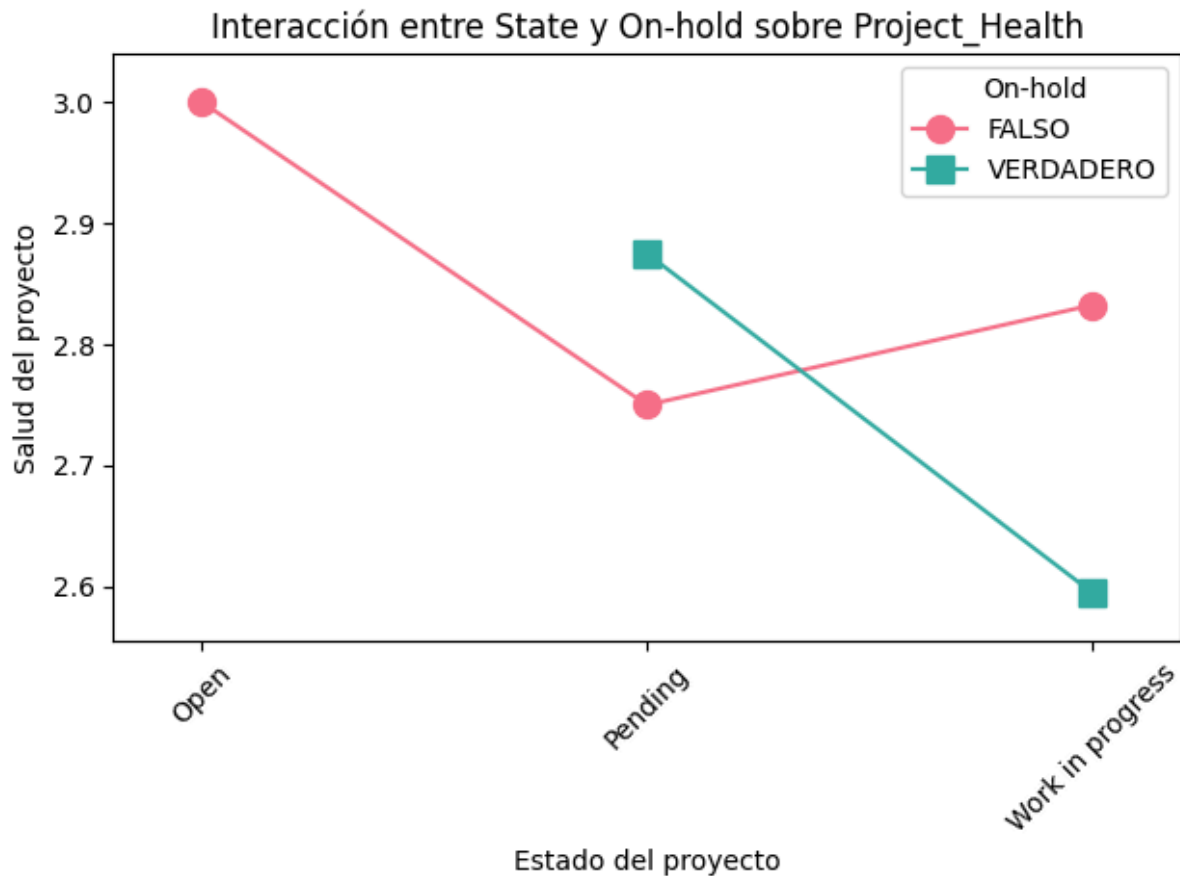
En general, la condición **On-hold = VERDADERO** se asocia con una **menor salud promedio**, especialmente en la categoría **Work in progress**, donde se observa un **descenso visible respecto a FALSO**.

La **variación por State es leve**: *Open* tiende a valores ligeramente mayores, mientras que *Pending* y *Work in progress* muestran medias similares.

Conclusión

El análisis confirma que **la condición On-hold afecta directamente la salud del proyecto**, reduciendo sus niveles promedio, pero **sin modificar el patrón general entre estados**.

La **ausencia de interacción estadística** respalda que el **efecto negativo de estar en pausa es consistente** en todos los estados del proyecto.



Media de *Project_Type* por *Project_size*

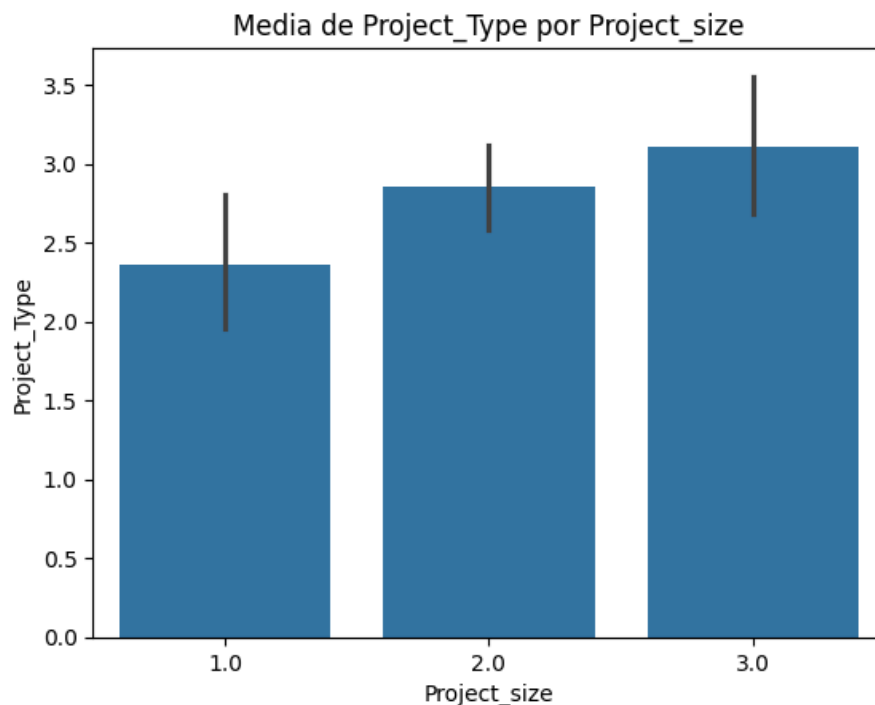
El gráfico de barras muestra la **media de *Project_Type*** agrupada según el **tamaño del proyecto (*Project_size*)**.

Se observa una tendencia ascendente: a medida que aumenta el tamaño del proyecto, también lo hace el valor promedio de *Project_Type*.

- Los **proyectos pequeños (1)** tienden a concentrarse en tipos con menor complejidad o alcance.
- Los **proyectos medianos (2)** presentan una media intermedia, reflejando una mayor variedad de categorías.
- Los **proyectos grandes (3)** alcanzan los valores promedio más altos, lo que sugiere que este grupo está asociado con tipos de proyecto más avanzados, estratégicos o de infraestructura.

Las **barras de error (IC 95%)** se superponen parcialmente, lo que indica que aunque existe una **tendencia positiva clara**, las diferencias entre grupos podrían no ser completamente significativas a nivel estadístico.

Sin embargo, la relación visual sugiere que el **tipo de proyecto aumenta en complejidad conforme crece el tamaño del mismo**, lo cual resulta coherente con la estructura organizacional de FORVIA, donde los proyectos de gran tamaño suelen involucrar más recursos, fases técnicas y coordinación transversal.



Análisis de varianza (ANOVA) y MANOVA para *Project_Type*

El objetivo de este análisis fue determinar si existen **diferencias significativas en el tipo de proyecto (Project_Type)** en función del **tamaño del proyecto (Project_size)** y del **estado operativo (State)**.

$$\text{Project_Type} = \mu + \text{Project_size} + \varepsilon$$

Fuente de variación	gl	Suma de cuadrados	Media cuadrática	F	p-valor
Project_size	2	20.39	10.19	3.53	0.0307
Residual	243	701.01	2.88	—	—

El **p-valor = 0.0307 (< 0.05)** indica que existen **diferencias estadísticamente significativas** entre los grupos de tamaño de proyecto.

Esto significa que el **tipo de proyecto varía según la magnitud o escala del mismo**, confirmando que los proyectos más grandes tienden a estar asociados con categorías más complejas o estratégicas (*Project_Type* más alto).

Este comportamiento se observa también en la gráfica de barras, donde la media de *Project_Type* aumenta progresivamente de los proyectos pequeños (1) hacia los grandes (3).

MANOVA simplificado: efectos combinados de *Project_size* y *State*

$$\text{Project_Type} = \mu + \text{Project_size} + \text{State} + (\text{Project_size} \times \text{State}) + \varepsilon$$

Fuente	gl	Suma de cuadrados	Media cuadrática	F	p-valor
Project_size	2	20.39	10.19	3.55	0.0302
State	2	13.26	6.63	2.31	0.1013

Project_size × State	4	5.13	1.28	0.45	0.7746
Residual	239	685.73	2.87	—	—

Los resultados muestran que:

- El **tamaño del proyecto (*Project_size*)** tiene un **efecto significativo ($p < 0.05$)** sobre el tipo de proyecto.
- El **estado (*State*)** no presenta un efecto significativo ($p > 0.05$).
- La **interacción entre ambos factores** tampoco es significativa ($p = 0.7746$), lo que sugiere que el efecto del tamaño es consistente en todos los estados.

Interpretación general

Los resultados del ANOVA y MANOVA confirman que el **tipo de proyecto depende principalmente del tamaño del proyecto**, mientras que el estado operativo no modifica de forma significativa esta relación.

En términos prácticos:

- Los **proyectos grandes ($size = 3$)** se asocian con tipos más avanzados (*Project_Type* altos), generalmente relacionados con infraestructura o coordinación global.
- Los **proyectos pequeños ($size = 1$)** corresponden con tipos más simples o de menor alcance técnico.
- No se evidenció una interacción significativa entre el tamaño y el estado, lo que indica que esta relación se mantiene estable independientemente de la etapa del proyecto (*Open, Pending, Work in progress*).

En conclusión, el **Project_size es un factor determinante** en la clasificación y complejidad del tipo de proyecto, validando su uso como variable predictora clave en el análisis de portafolios de FORVIA.

Distribución de *Project_Type* por *State*

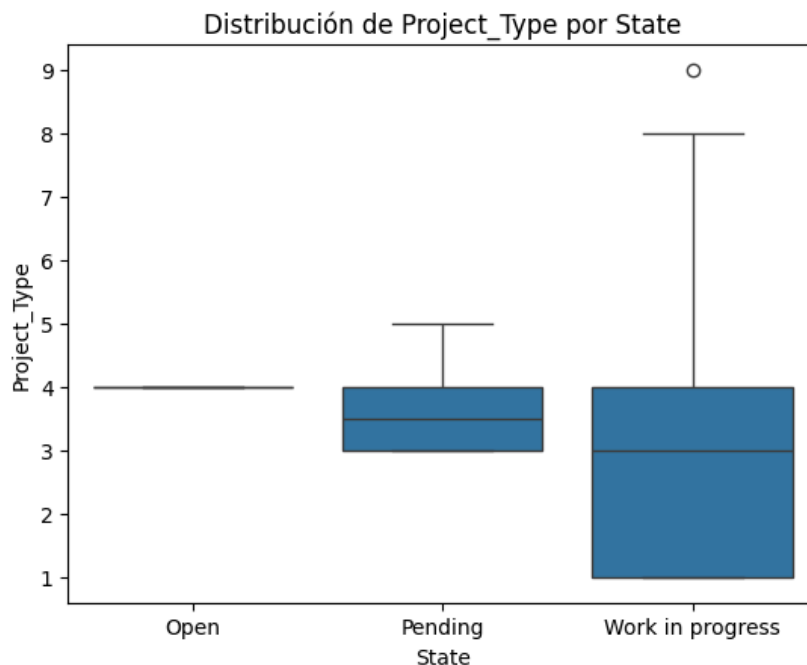
El **gráfico tipo boxplot** representa la **distribución del tipo de proyecto (*Project_Type*)** en función del **estado operativo (*State*)**.

Se observa que los proyectos “**Work in progress**” presentan la **mayor dispersión en los valores de *Project_Type***, abarcando desde categorías simples hasta tipos más complejos, con algunos **valores atípicos cercanos a 9**.

Por otro lado, los proyectos “**Pending**” muestran una **distribución más concentrada alrededor de valores medios** (entre 3 y 4), mientras que los “**Open**” presentan valores **constantes en torno a 4**, reflejando **menor variabilidad**.

Estos resultados coinciden con el **ANOVA de dos factores aplicado**, en el cual la variable **State** no mostró un efecto estadísticamente significativo ($p = 0.1013 > 0.05$).

En consecuencia, se concluye que el **tipo de proyecto (*Project_Type*) no depende directamente del estado operativo**, aunque los proyectos en curso (*Work in progress*) tienden a exhibir **mayor diversidad de tipos**, lo que refleja **diferentes niveles de avance y complejidad** dentro del portafolio de proyectos de FORVIA.



Interacción *Project_size* × *State* sobre *Project_Type*

La figura muestra el **interaction plot** de las medias de *Project_Type* según el tamaño del proyecto (*Project_size* = 1, 2, 3), diferenciando por el estado operativo (*State*: Open, Pending, Work in progress).

Lecturas visuales

Se observa una **tendencia ascendente** de *Project_Type* conforme aumenta *Project_size*, especialmente clara en la categoría **Work in progress**.

Las **líneas entre estados son casi paralelas**, lo que sugiere **ausencia de interacción significativa**: el efecto del tamaño sobre el tipo de proyecto se mantiene **consistente** en *Open*, *Pending* y *Work in progress*.

Asimismo, el estado **Pending** exhibe niveles ligeramente más altos al pasar a **Large (3)**, mientras que en **Work in progress** el crecimiento con el tamaño es más gradual.

Contraste estadístico (ANOVA de dos factores)

A partir del modelo

Project_Type ~ *Project_size* + *State* + *Project_size*:*State*

se obtuvieron los siguientes resultados:

Factor	F	p-valor	Significancia
<i>Project_size</i>	3.55	0.0302	Significativo
<i>State</i>	2.31	0.1013	No significativo
<i>Project_size</i> × <i>State</i>	0.45	0.7746	Sin interacción

Conclusión

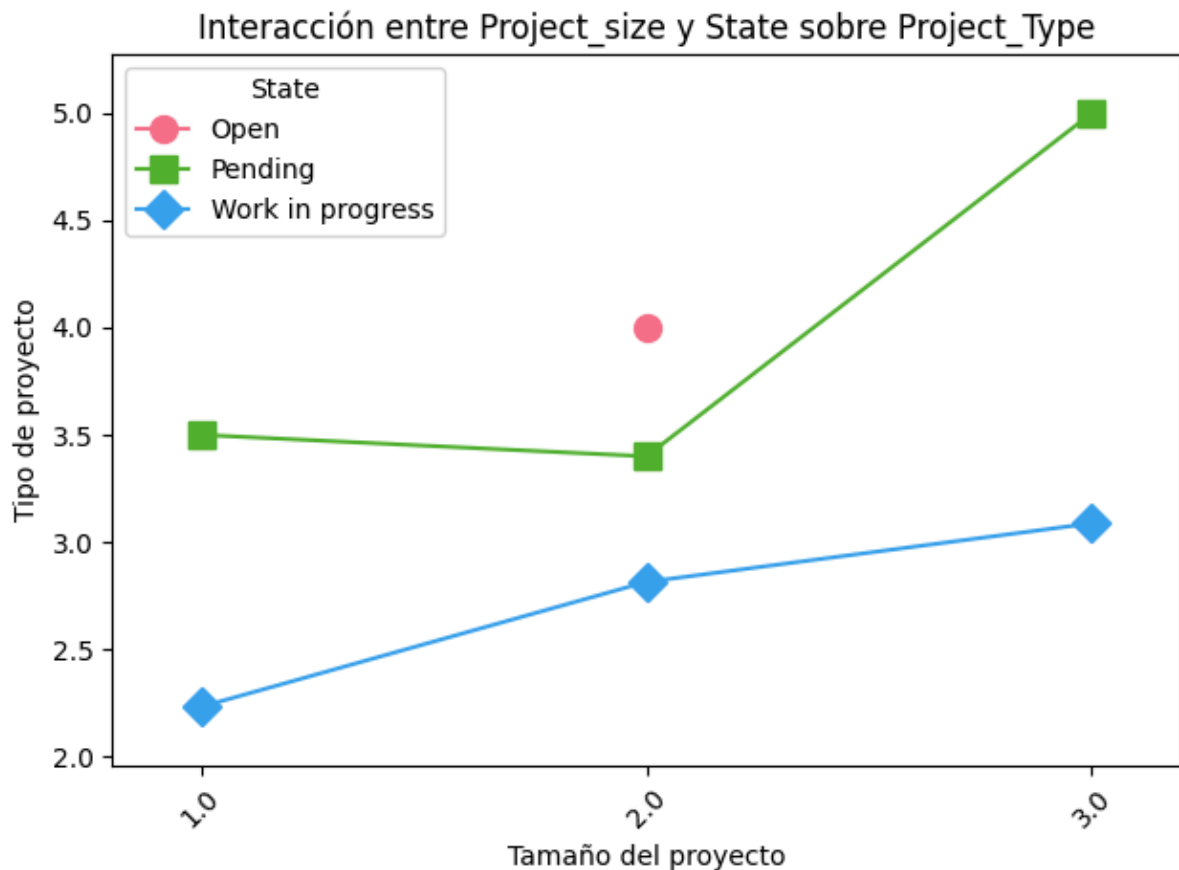
El **tipo de proyecto (Project_Type)** tiende a **incrementar con el tamaño del proyecto (Project_size)**, de forma independiente al estado operativo.

El **estado (State)** no altera significativamente esta relación, ni la modula de manera conjunta, lo que confirma la **ausencia de interacción estadística**.

Implicación práctica

Para **anticipar la complejidad o categoría del proyecto**, el **tamaño (Project_size)** es el **predicor principal**.

El estado operativo (*Open*, *Pending*, *Work in progress*) aporta información secundaria, pero **no modifica el patrón general** del tipo de proyecto.



Conclusiones

En conclusión, el análisis estadístico aplicado mediante **ANOVA y ANOVA de dos factores** permitió identificar los **principales factores que explican las diferencias en el desempeño de los proyectos de FORVIA**, considerando variables clave como el porcentaje de avance (*Percent_complete*), el tamaño del proyecto (*Project_size*), la salud (*Project_Health*) y el tipo de proyecto (*Project_Type*). Los resultados evidencian que el **estado operativo (State)** es el factor más influyente: los proyectos en “**Work in progress**” alcanzan niveles de avance significativamente superiores, lo que demuestra la relación directa entre la etapa de ejecución y el progreso del proyecto. Asimismo, el **grupo de negocio (BG)** mostró un efecto estadísticamente significativo sobre el tamaño, destacando a las divisiones **SAS** y **GROUP** como responsables de los proyectos de mayor magnitud, reflejando diferencias estructurales en alcance y recursos entre áreas.

Por su parte, la variable **On-hold** evidenció un impacto negativo sobre la salud del proyecto, indicando que las pausas o interrupciones reducen su estabilidad y desempeño general. En contraste, el **tipo de proyecto (Project_Type)** y las **interacciones entre factores** no mostraron efectos significativos, lo que sugiere que los patrones de comportamiento del portafolio son consistentes entre categorías. En conjunto, estos hallazgos subrayan la importancia de **priorizar la ejecución continua de los proyectos, optimizar la planificación por grupo de negocio y mantener la operación activa** como estrategias clave para fortalecer la eficiencia y la salud general del portafolio de proyectos en FORVIA.