

1. Introducción

Este análisis tiene como objetivo explorar y caracterizar la base de datos proyectos_forvia.csv, de Forvia, con el fin de identificar patrones, distribuciones y tendencias en variables categóricas y numéricas relacionadas con los proyectos.

2. Metodología

Fuente de datos: proyectos_forvia.csv

Limpieza de datos:

- Eliminación de registros y columnas irrelevantes.
- Conversión de porcentajes a valores numéricos.
- Corrección y estandarización de fechas.
- Imputación de valores nulos (relleno con *backfill* y reemplazo con "NA").

Herramientas utilizadas: Python (Pandas, NumPy, Matplotlib).

Análisis aplicado:

- Tablas de frecuencia de variables categóricas.
- Gráficos de barras para distribución visual.
- Categorización de variables numéricas mediante la regla de Sturges.

3. Hallazgos Principales

3.1 Limpieza de datos

- Valores nulos iniciales:

Number	0
Active	0
Project Name	0
Project Type	0
Geographical scope	0
Project manager	0
State	0
Percent complete	0
Project size	0
Project organization	0
Planned Go Live date	0
Project target phase	173
Actual Go Live date	197
Actual end date	245
BG	0
Closed	244
Domain	0
Domain Path	0
Last WAR	13
Project Health	13
Project type	0
Recurrent activity	0
Planned start date	0
Actual start date	18
On-hold	0
dtype:	int64

- **Transformaciones realizadas:**
 - Porcentajes convertidos a valores numéricos.
 - Fechas estandarizadas.
- **Estado final del DataFrame:**

```

Number          0
Active          0
Project Name     0
Project Type     0
Geographical scope 0
Project manager  0
State           0
Percent complete 0
Project size     0
Project organization 0
Planned Go Live date 0
Project target phase 0
BG              0
Domain          0
Last WAR        0
Project Health   0
Project type     0
Recurrent activity 0
Planned start date 0
Actual start date 0
On-hold         0
dtype: int64

```

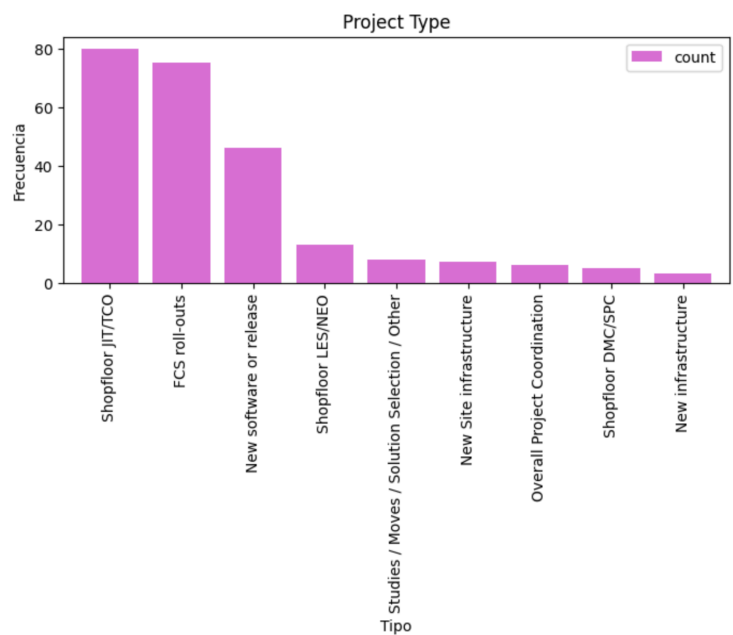
3.2 Variables Categóricas

Project Type

- Tabla de frecuencias:

	count
Project Type	
Shopfloor JIT/TCO	80
FCS roll-outs	75
New software or release	46
Shopfloor LES/NEO	13
Studies / Moves / Solution Selection / Other	8
New Site infrastructure	7
Overall Project Coordination	6
Shopfloor DMC/SPC	5
New infrastructure	3

● Gráfico:



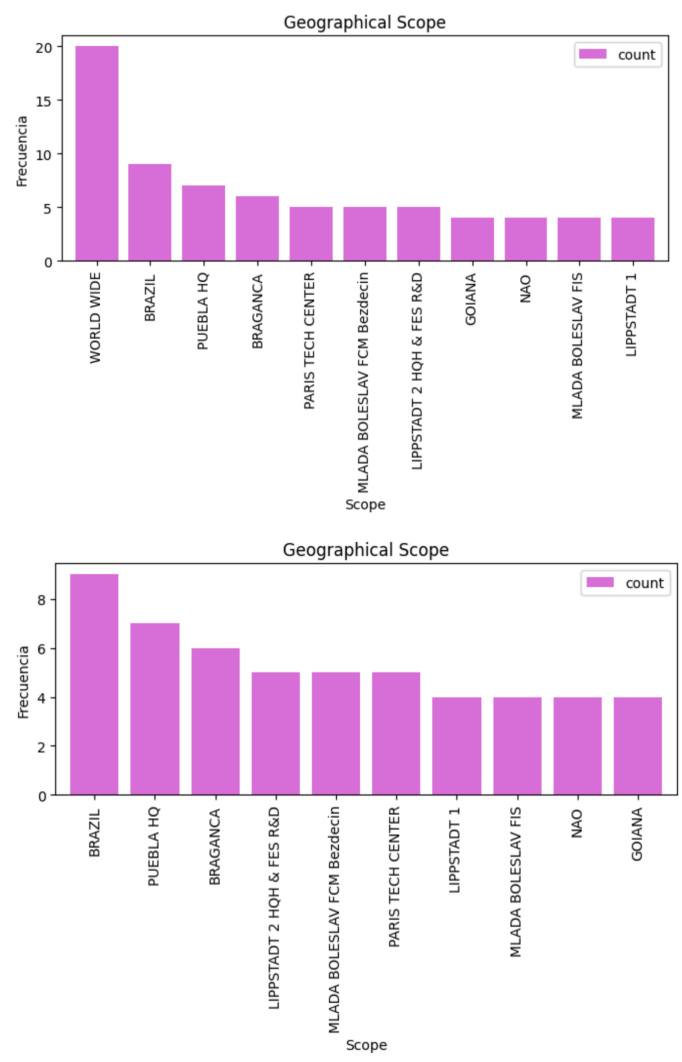
La distribución muestra que la operación o el ecosistema de proyectos se enfocan en tareas de "Shopfloor JIT/TCO" y "FCS roll-outs", mientras que las actividades de infraestructura, coordinación o estudios detallados representan una porción muy pequeña de las operaciones registradas.

Geographical Scope

● Tabla de frecuencias:

	count
Geographical scope	
WORLD WIDE	20
BRAZIL	9
PUEBLA HQ	7
BRAGANCA	6
PARIS TECH CENTER	5
MLADA BOLESLAV FCM Bezdecin	5
LIPPSTADT 2 HQH & FES R&D	5
GOIANA	4
NAO	4
MLADA BOLESLAV FIS	4
LIPPSTADT 1	4

● Gráfico:



La mayoría de las operaciones son globales, favoreciendo a la etiqueta "WORLD WIDE".

Las ubicaciones específicas a nivel de ciudad o centro técnico son mucho menos frecuentes y representan los valores marginales en la distribución.

Project Manager

- Tabla de frecuencias:

	count
Project manager	
SOUZA Alexandre	15
BENAVENT Pablo	12
LOUBET Diana	9
SVARICEK Dan	9
MALTER Karsten	7
CAEIRO Helder	6
LOURENCO Douglas	6
FENETRE Teddy	5
NITSCKIE Jade	5
SANZ Ana	4
FRENZ Ingo	4
ATTOF Mohamed	4
CUCHEROUSSET Christelle	4

- Gráfico:



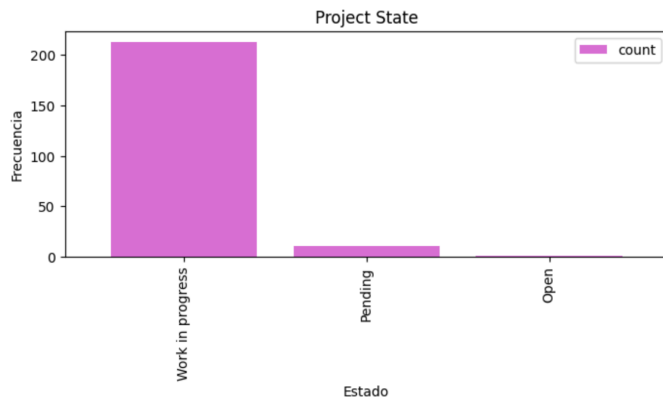
El patrón de actividad de los gerentes de proyecto tiene una distribución muy desigual, donde la mayor parte de la gestión de proyectos está concentrada en unos pocos individuos clave, con mayor gerentes con una participación menor y uniforme.

State

- Tabla de frecuencias:

	count
State	
Work in progress	213
Pending	11
Open	1

- Gráfico:



El estado del proyecto es abrumadoramente **"Work in progress"**, lo que indica que una vez que un proyecto es aprobado o iniciado, se convierte inmediatamente en un elemento de trabajo activo, el estado "Pending" se encuentra en 11, y el estado "Open" se encuentra en 1.

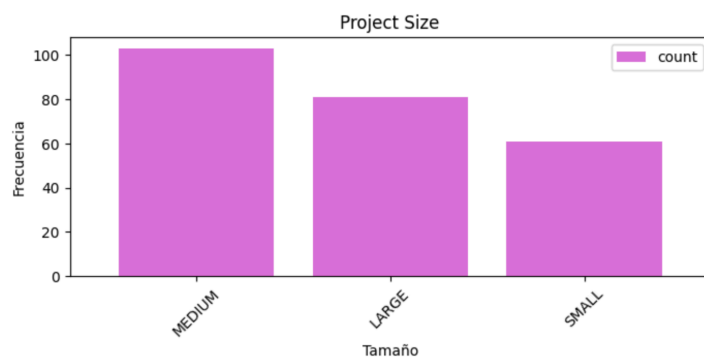
La gráfica demuestra una clara priorización y concentración de los elementos en ejecución ("Work in progress"), lo que implica una gestión muy activa.

Project Size

- Tabla de frecuencias:

Project size	count
MEDIUM	103
LARGE	81
SMALL	61

- Gráfico:



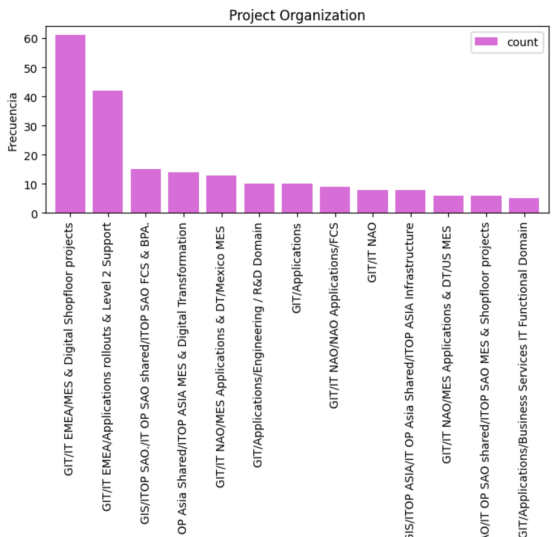
El patrón muestra que la mayoría de los proyectos registrados son de tamaño mediano y grande, lo que podría implicar que la organización se enfoca principalmente en proyectos con una escala de esfuerzo alto, mientras que los proyectos pequeños son minoría.

Project Organization

- Tabla de frecuencias:

Project organization	count
GIT/IT EMEA/MES & Digital Shopfloor projects	61
GIT/IT EMEA/Applications rollouts & Level 2 Support	42
GIS/ITOP SAO./IT OP SAO shared/ITOP SAO FCS & BPA.	15
GIS/ITOP ASIA/IT OP Asia Shared/ITOP ASIA MES & Digital Transformation	14
GIT/IT NAO/MES Applications & DT/Mexico MES	13
GIT/Applications/Engineering / R&D Domain	10
GIT/Applications	10
GIT/IT NAO/NAO Applications/FCS	9
GIT/IT NAO	8
GIS/ITOP ASIA/IT OP Asia Shared/ITOP ASIA Infrastructure	8
GIT/IT NAO/MES Applications & DT/US MES	6
GIS/ITOP SAO/IT OP SAO shared/ITOP SAO MES & Shopfloor projects	6
GIT/Applications/Business Services IT Functional Domain	5

- Gráfico:



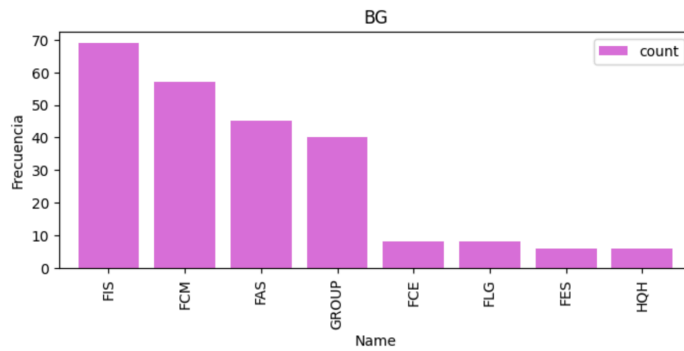
La distribución muestra un patrón desigual donde la gestión de proyectos está fuertemente centralizada o es más productiva dentro de la región GTI EMA..., con todas las demás organizaciones operando a un nivel de frecuencia significativamente inferior.

BG

- Tabla de frecuencias:

	count
BG	
FIS	69
FCM	57
FAS	45
GROUP	40
FCE	8
FLG	8
FES	6
HQH	6

- Gráfico:



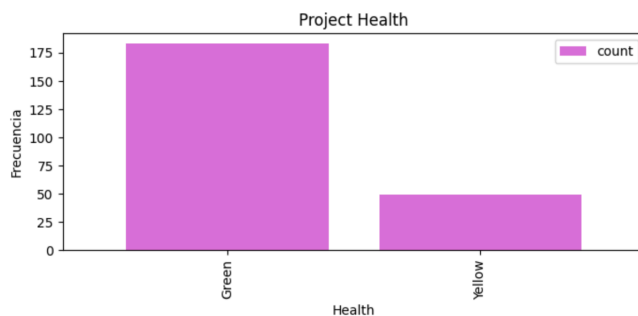
El patrón muestra una clara dependencia de los grupos de BG, FIS, FCM, FAS y GROUP para las operaciones registradas. Las otras categorías de grupos de proyectos son marginales en volumen, lo que indica un foco operativo muy concentrado en los cuatro principales grupos.

Project Health

- Tabla de frecuencias:

	count
Project Health	
Green	183
Yellow	49

- Gráfico:



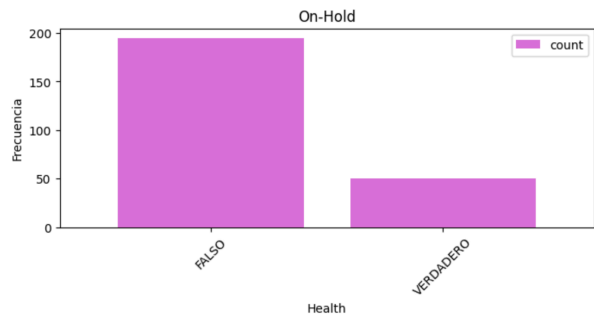
El patrón muestra que el ecosistema de proyectos son saludables, con casi cuatro de cada cinco proyectos reportados como "Green".

On-Hold

- Tabla de frecuencias:

	count
On-hold	
FALSO	195
VERDADERO	50

- Gráfico:



El patrón muestra que el estado de retención ("On-Hold") es infrecuente, con una tendencia a mantener los proyectos o elementos en un estado activo.

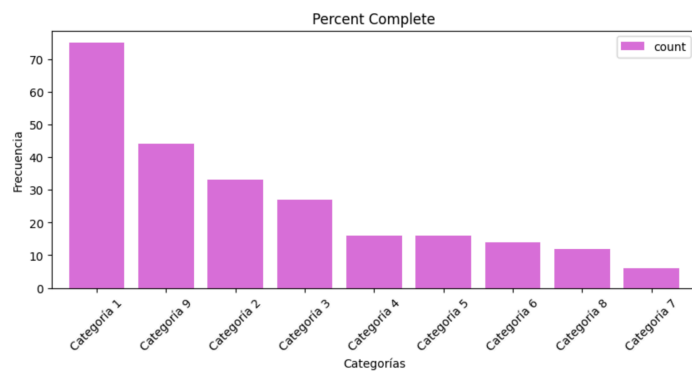
3.3 Variable Numérica

Percent Complete (Regla de Sturges)

- Tabla de categorías:

	count
Percent complete	
Categoría 1	75
Categoría 9	44
Categoría 2	33
Categoría 3	27
Categoría 4	16
Categoría 5	16
Categoría 6	14
Categoría 8	12
Categoría 7	6

- Gráfico de barras con frecuencias:



El patrón muestra que el proceso está muy activo en la creación de elementos (Categoría 1) y que un alto número de elementos está cerca la (Categoría 9), lo que implica una eficiencia variable.

4. Conclusión

El análisis de la base de datos muestra que Forvia concentra sus esfuerzos en proyectos de mediana y gran escala, principalmente en áreas clave como "Shopfloor JIT/TCO" y "FCS roll-outs", así como en los grupos GROUP, FIS, FCM y FAS. La mayoría de los proyectos se encuentran activos y en buen estado "Green", con pocas interrupciones, lo que refleja estabilidad y eficiencia, aunque se observa una fuerte centralización en pocos gerentes y un uso generalizado de la etiqueta "WORLD WIDE", lo que limita la precisión en el seguimiento geográfico por lo que se eliminó. En general, los resultados demuestran un sistema operativo sólido, aunque sí existe una área de mejora en la distribución de responsabilidades y la clasificación de proyectos.