



Departamento Nacional de Planeación
www.dnp.gov.co

INVERSIÓN PARA EL CUMPLIMIENTO DE LOS ODS

Unidad de Científicos de Datos
Dirección de Desarrollo Digital

Julio, 2018
dnp.gov.co

- 1.Introducción
- 2.Metodología
- 3.Resultados
- 4.Conclusiones



INVERSIÓN PARA EL CUMPLIMIENTO DE LOS ODS

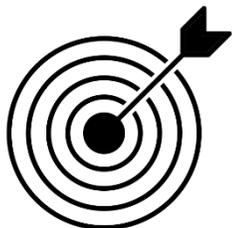
Introducción

1

INVERSIÓN PARA EL CUMPLIMIENTO DE LOS ODS

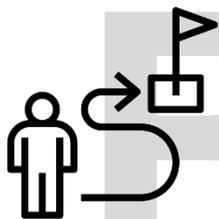
Inversión para el cumplimiento de los ODS

Clasificación de documentos con inversión económica en los 16 ODS



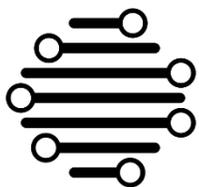
Objetivo

Clasificar diferentes documentos que representan inversiones económicas, según su correspondencia con cada ODS



Metodología

Vectorización de los documentos según la información disponible de los Objetivos de Desarrollo Sostenible para usar una medida de similitud entre éstos.



Insumos

24379 Proyectos de inversión de SGR, PGN y SUIFT Territorial

14731 Documentos de funcionamiento de hacienda y cooperación internacional

7580 Documentos de proyectos privados de Prosperidad Social (Mapa Social)

86 Documentos de SGP y Entidad Territorial Descentralizada

Para un total de **46776** textos

Clasificación gasto en ODS

Construcción de una herramienta de rastreo y seguimiento de la financiación de la Agenda 2030

- Seguimiento a la financiación de los Objetivos de Desarrollo Sostenible.
- Identificación fuentes de información
- Revisión y análisis de documentos relacionados
- Consolidar conjunto de palabras clave por objetivo y meta
- Diseño metodología de minería de texto
- Diseño, prueba y evaluación algoritmo de clasificación



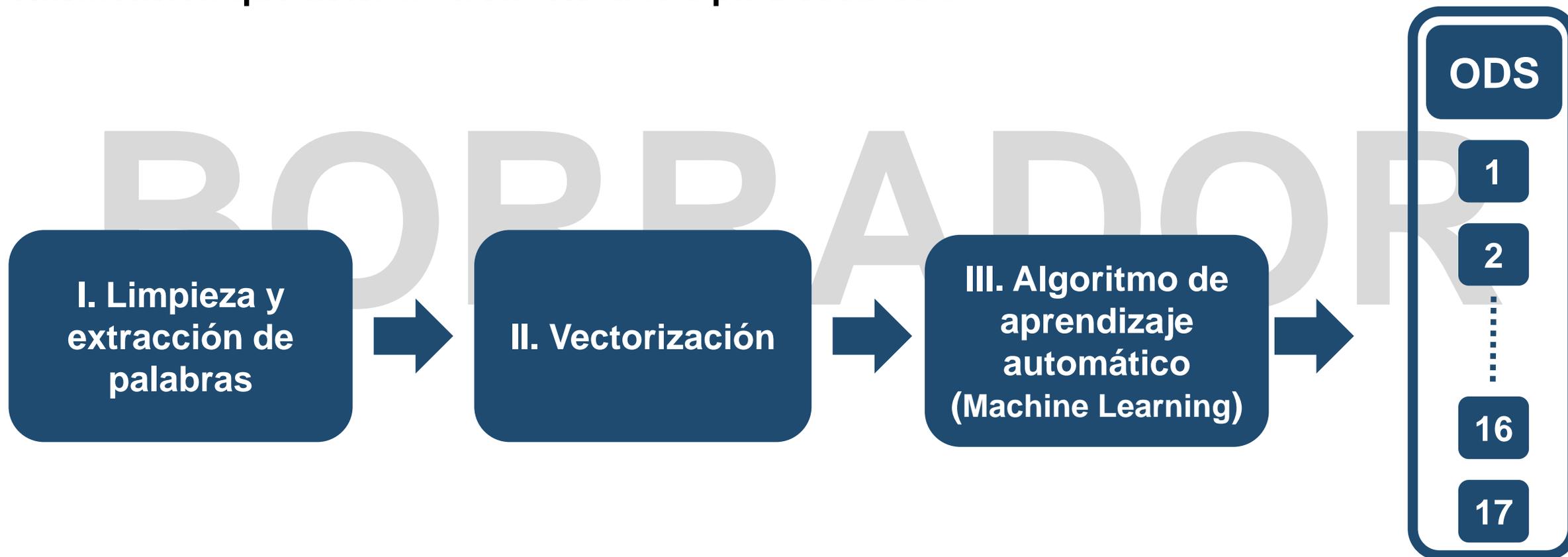
1. Metodología



INVERSIÓN PARA EL CUMPLIMIENTO DE LOS ODS

Análisis de texto para correspondencia con los ODS

Sobre el texto contenido en los documentos se realizan procesos de limpieza y extracción de palabras relevantes. El texto obtenido se representa numéricamente tomando en cuenta su estructura. Esta representación numérica de los documentos es suministrada a un módulo de clasificación que determina su relevancia para cada ODS



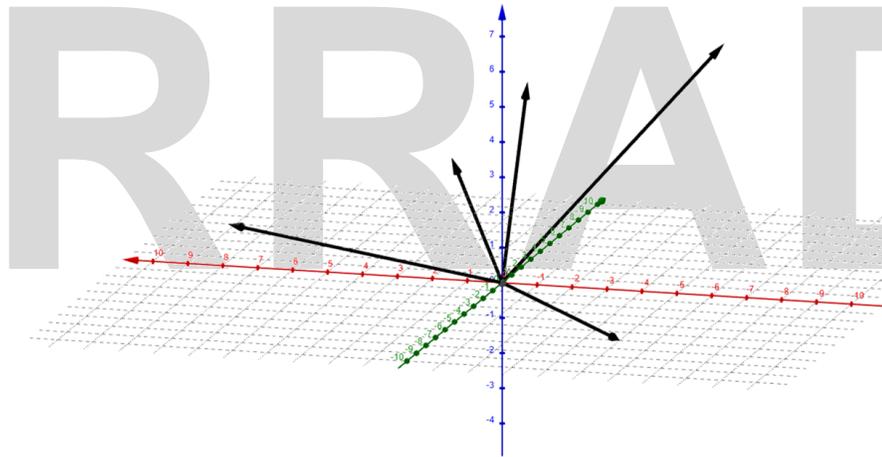
Caracterización a partir del texto ODS

La metodología implementada se basa en el análisis del texto que define cada ODS, sus metas e indicadores con el fin de construir una representación vectorial en donde es posible utilizar medidas de similitud

Palabras



Vectores



Complicaciones:

- Poco texto exclusivo a cada ODS
- Las categorías **no** son mutuamente excluyentes
- Caracterización de un categoría “18”

Vectorización del texto

Extracción de información en forma cuantitativa

	Palabra 1	Palabra 2	Palabra 3	Palabra 4	Palabra 5	Palabra 6	Palabra 7	Palabra 8
Obj 1	0,2	0	0,4	0,3	0	0,1	0	0,2
Obj 2	0	0,2	0,4	0	0,2	0,3	0	0
Obj 3	0,4	0	0,1	0,3	0	0,1	0	0,1
Obj 4	0	0,1	0	0,2	0	0	0,1	0
Obj 5	0	0	0,2	0	0	0,4	0	0
Obj 6	0,1	0,1	0	0,2	0	0,1	0,1	0,3
Obj 7	0,2	0,1	0,3	0,4	0	0,2	0	0,2

- Remoción de palabras irrelevantes (*stopwords*).
- Peso de palabras en función de su ocurrencia y el número de documentos en los que aparecen

Características vectoriales de los ODS

Tendencias de agrupación de ODS

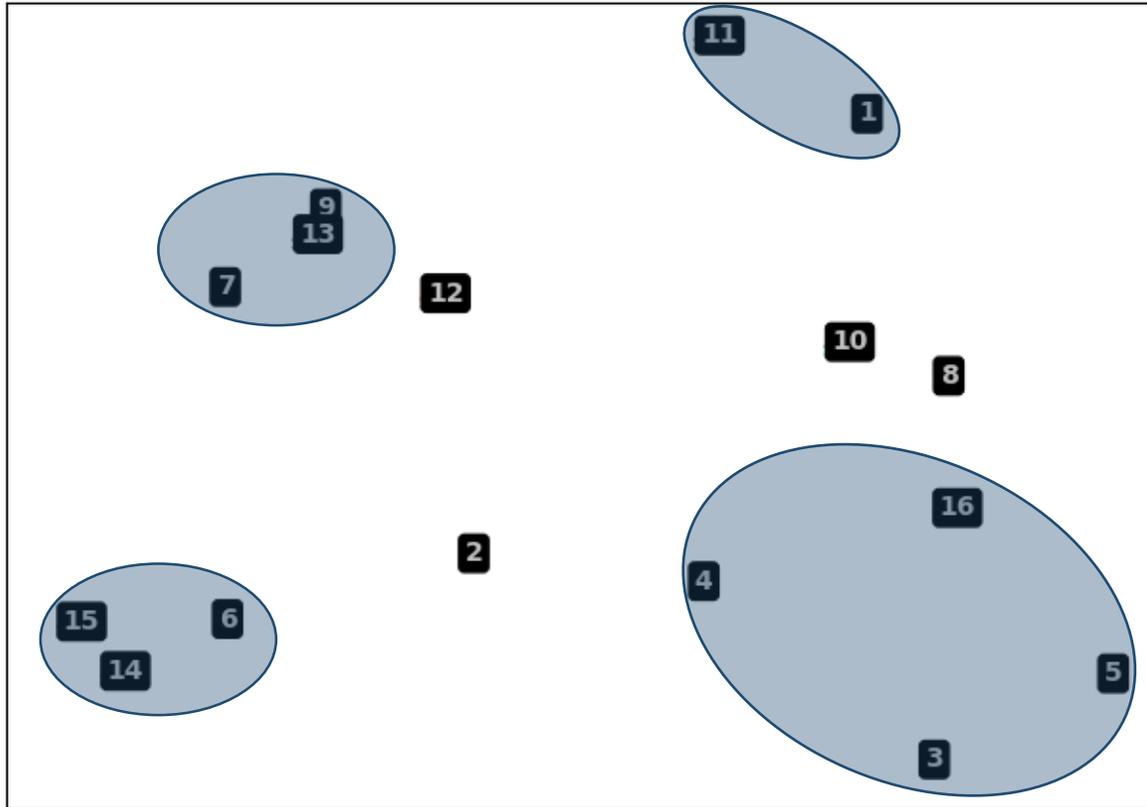


Gráfico 1. Reducción de dimensionalidad con método ACP*

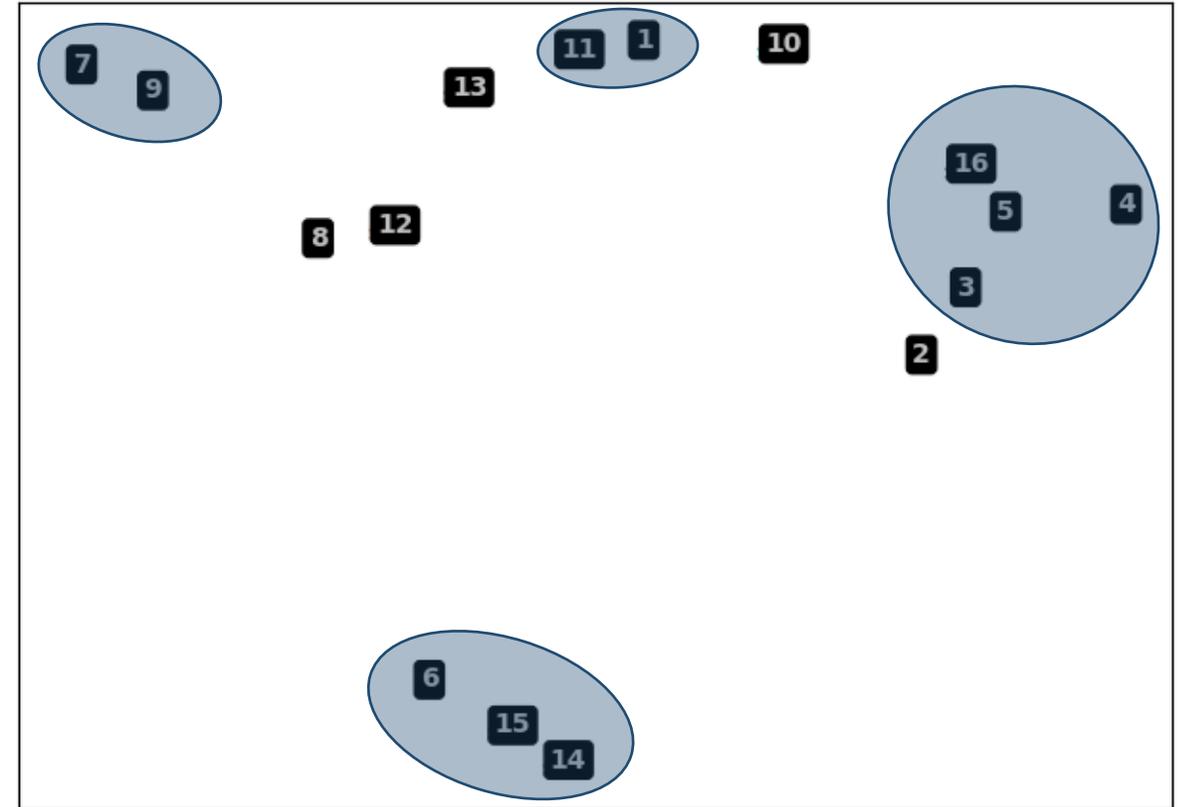


Gráfico 2. Reducción de dimensionalidad con método t-SNE**

*Análisis de Componentes Principales
**t-student Stochastic Neighbor Embedding

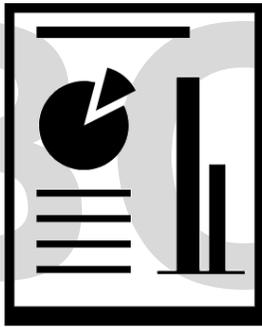
Preprocesamiento de proyectos

Sobre los documentos a clasificar se surte una etapa previa de preprocesamiento y transformación del texto utilizando la matriz de términos construida en función de los ODS



Fichas Estadísticas Básicas de Inversión - SGR

Resume las características centrales de un proyecto o programa. Esta ficha deber ser diligenciada por las entidades para cada uno de los proyectos o programas que requiera financiamiento



Del total de documentos disponibles en SUIFP-SGR las fichas EBI contienen la información necesaria para capturar de manera automatizada aspectos cualitativos y cuantitativos del proyecto como:

1. Nombre del proyecto
2. Descripción
3. Sector
4. Subsector

Clasificación gasto en ODS: Resultados experimentales SGR

Proyectos y ODS vectorizados, con reducción de dimensionalidad para ser visualizados

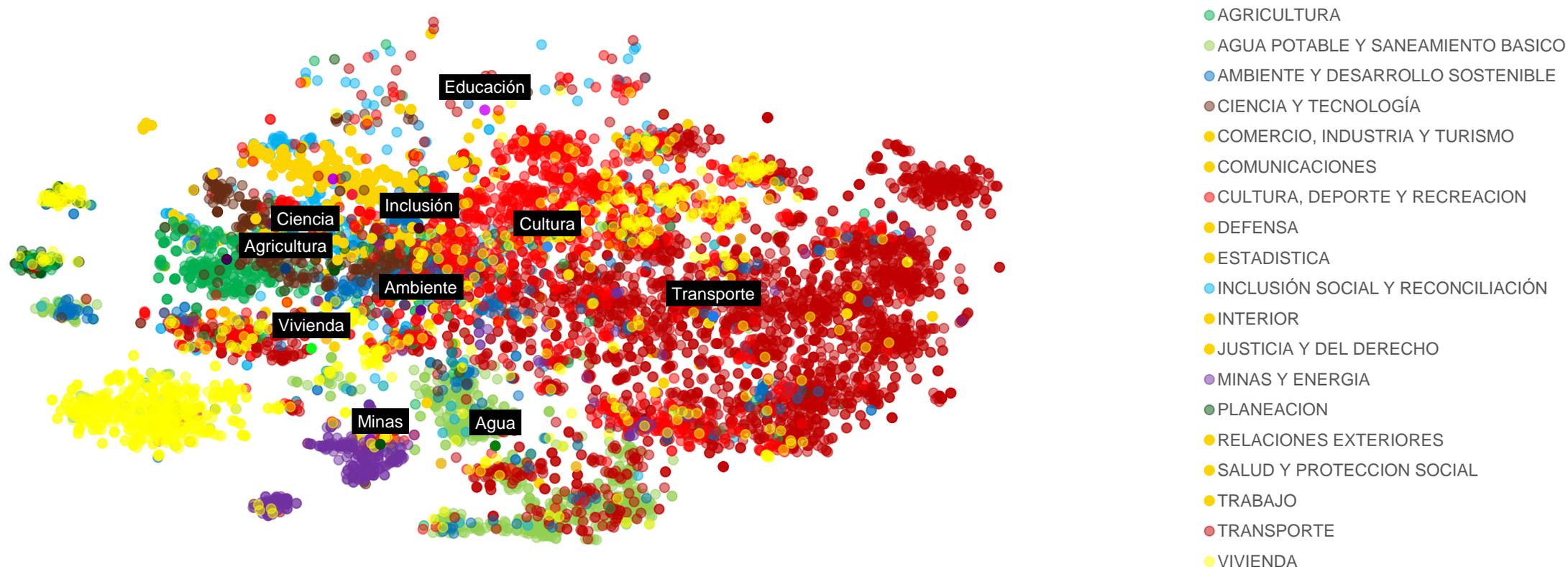


Gráfico 3. *t*-SNE de representaciones vectoriales de proyectos y ODS

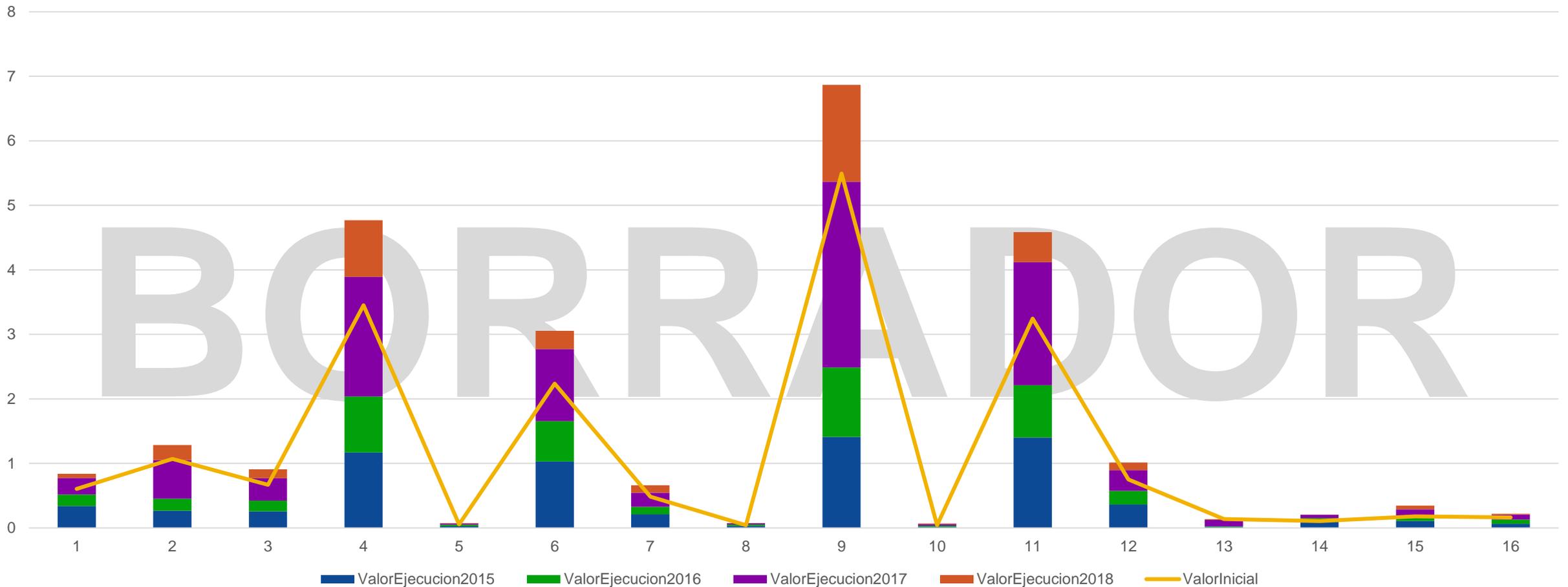
Resultados

3

INVERSIÓN PARA EL CUMPLIMIENTO DE LOS ODS

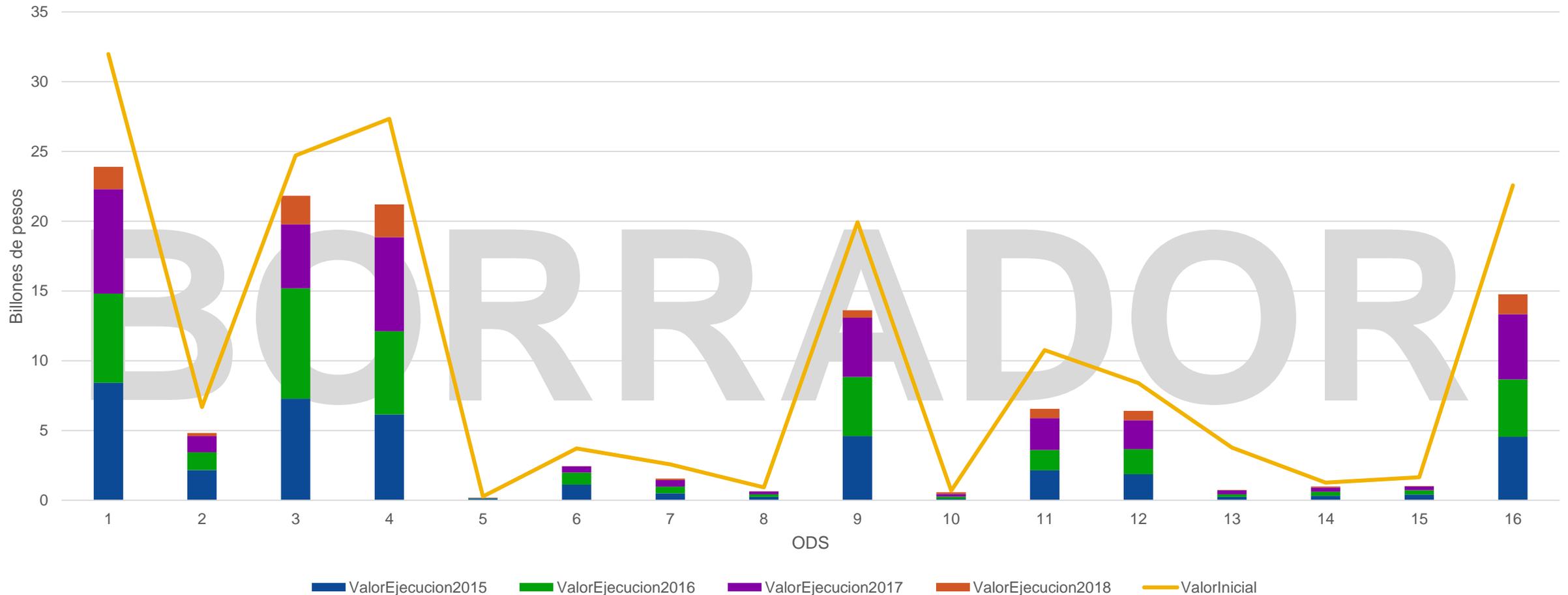
Inversión a ODS a través de proyectos SGR

Inversión por año de acuerdo con la clasificación de proyectos SGR en los ODS



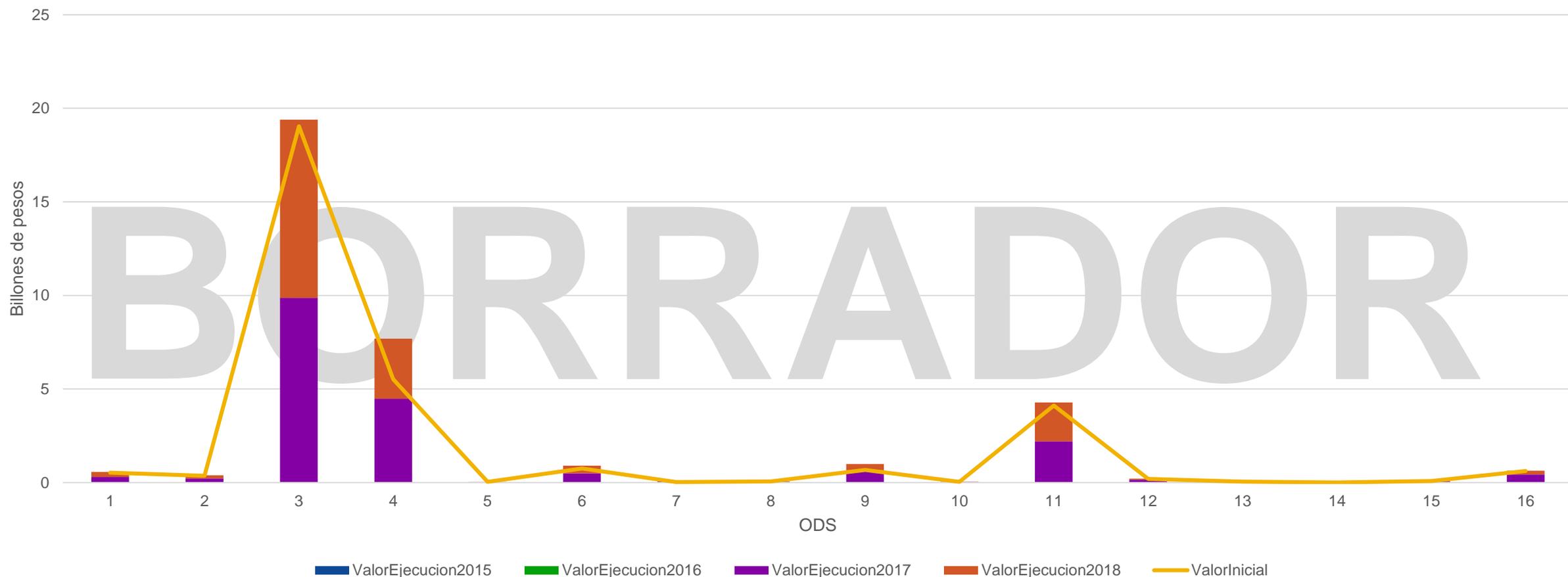
Inversión a ODS a través de proyectos PGN

Inversión por año de acuerdo con la clasificación de proyectos PGN en los ODS



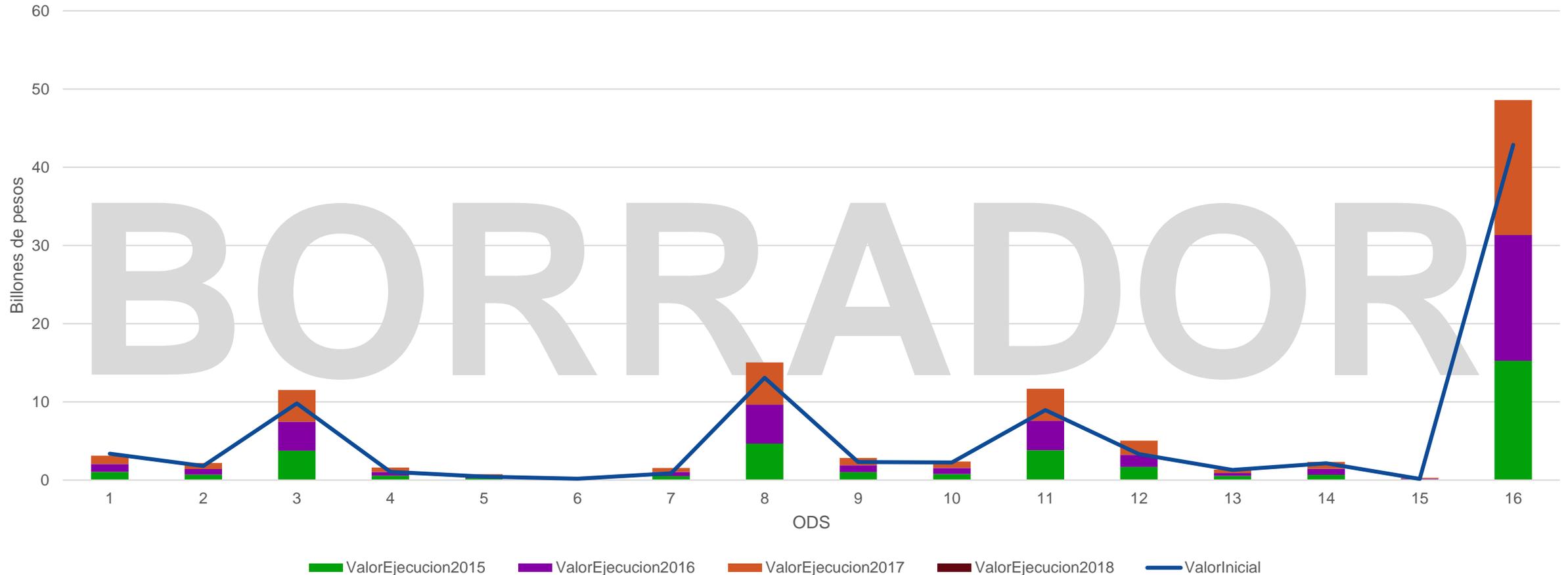
Inversión a ODS a través de proyectos SUIFP Territorial

Inversión por año de acuerdo con la clasificación de proyectos de SUIFP Territorial en los ODS



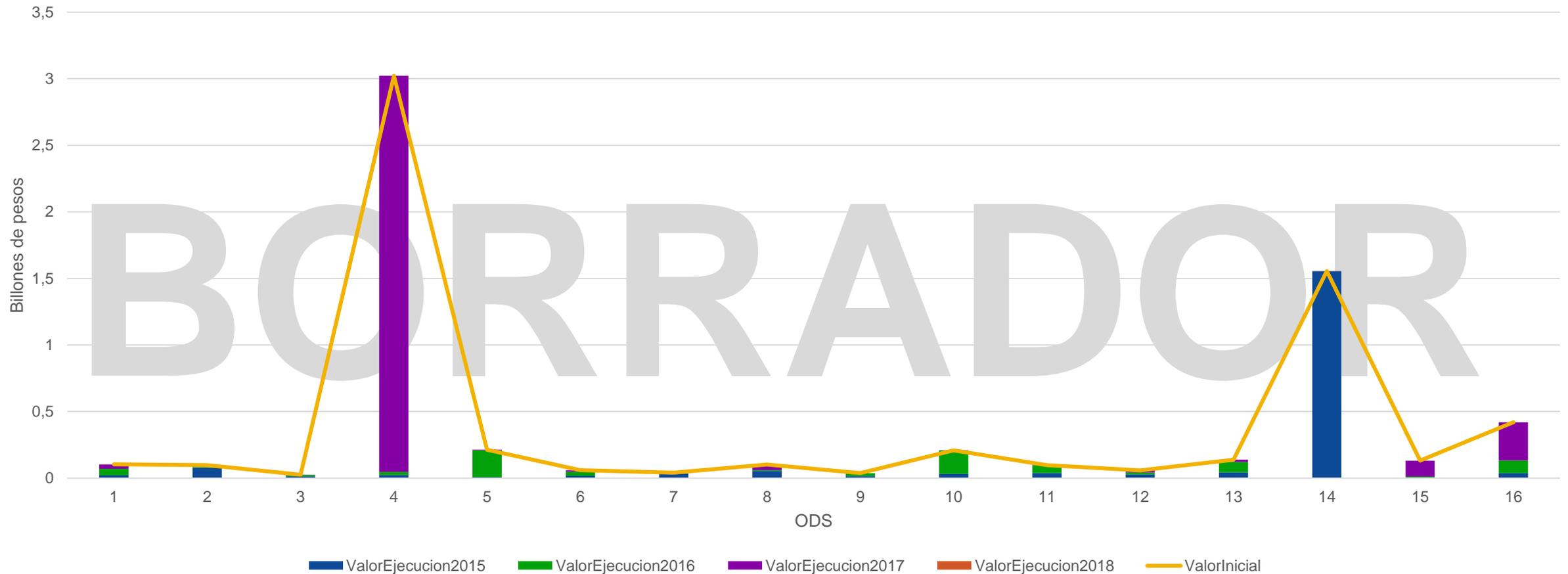
Inversión a ODS a través de funcionamiento de Hacienda

Inversión por año de acuerdo con la clasificación de los rubros de funcionamiento de hacienda en los ODS



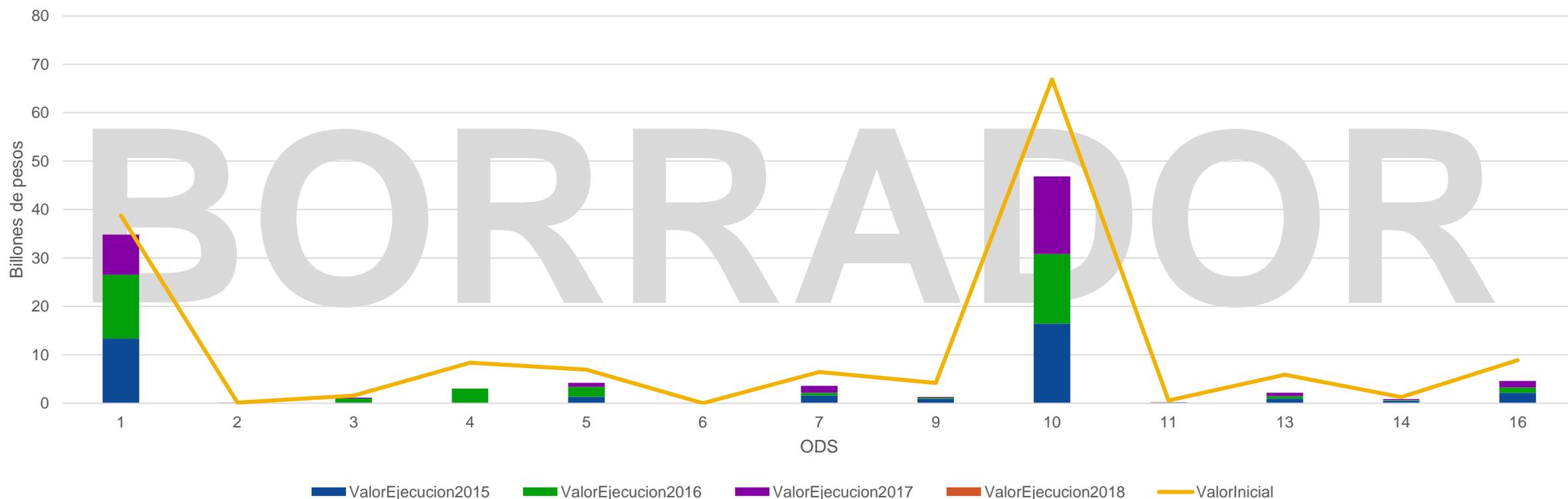
Inversión a ODS a través de cooperación internacional

Inversión por año de acuerdo con la clasificación de los rubros de cooperación internacional en los ODS



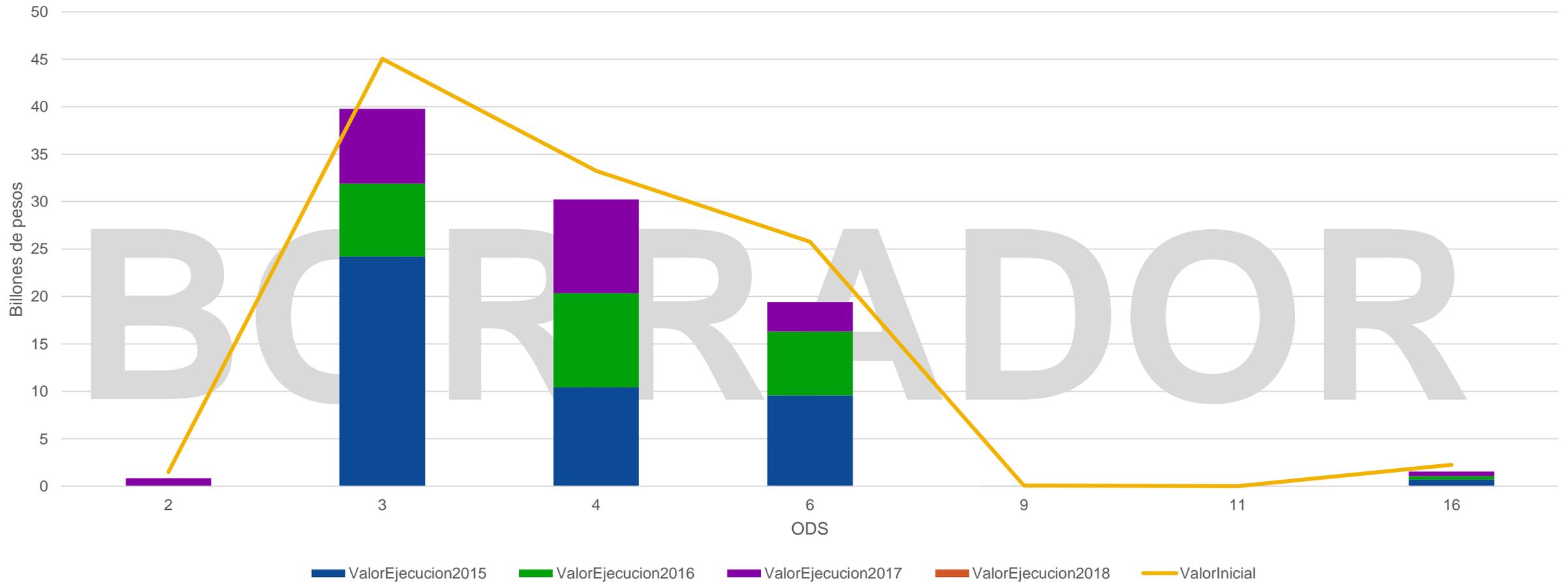
Inversión a ODS a través de entidad territorial descentralizada

Inversión por año de acuerdo con la clasificación de rubros de la entidad territorial descentralizada en los ODS



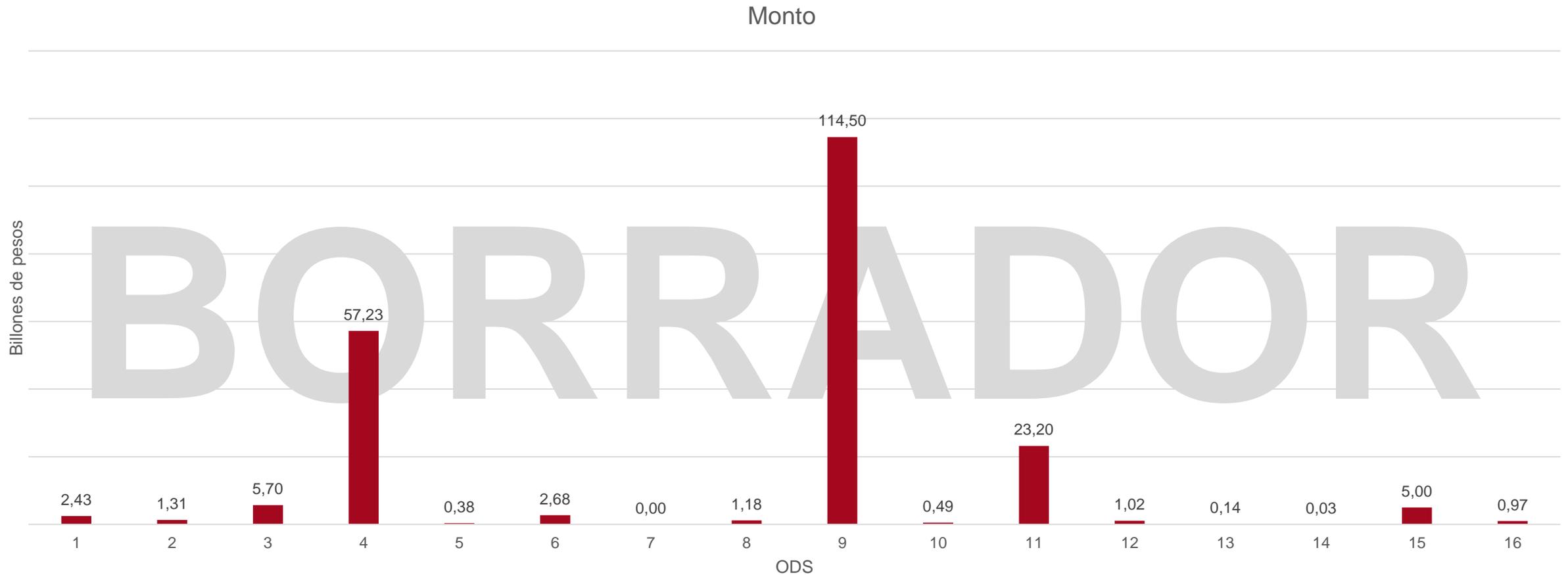
Inversión a ODS a través de SGP

Inversión por año de acuerdo con la clasificación SGP en los ODS



Inversión a ODS a través de Mapa Social

Monto por ODS de los proyectos privados de mapa social.



Conclusiones

4

INVERSIÓN PARA EL CUMPLIMIENTO DE LOS ODS

Conclusiones

- Como alternativa al problema que existe de poco texto disponible de los ODS, se tiene a la proyección de diferentes documentos CONPES en los que se busca adjuntar al texto pre-existente, nuevos documentos con temáticas relacionadas.
- Aún cuando las clasificaciones de los textos predominen en algún ODS, no necesariamente la inversión es directamente proporcional.
- Se tiene que para el modelo actual, el error de clasificación es del 40%.
- Bajo un conjunto de 600 documentos con clasificación previa, puede hacerse uso de metodologías más sofisticadas y precisas que la que actualmente se usa.
- La discusión de la clasificación de los rubros o documentos, queda abierta, según el enfoque que se quiera dar. Pues podría hacerse también una clasificación multietiqueta e incluso detallada a los tipos de actividades