

**SABER EXPO BI 5: UNA HERRAMIENTA DE INTELIGENCIA DE
NEGOCIOS WEB PARA LA GESTIÓN DE INDICADORES DE LAS
PRUEBAS SABER 5 EN LAS INSTITUCIONES EDUCATIVAS DE LA
SUBREGIÓN DE OBANDO DEL DEPARTAMENTO DE NARIÑO**

MANUAL DE REFERENCIA



INTRODUCCIÓN

En este documento se describe la documentación técnica de programación, incluyendo la instalación del entorno de desarrollo, la estructura de la aplicación, su compilación, la configuración de los diferentes servicios de integración utilizados.

METODOLOGIA DE DESARROLLO

La investigación fue de tipo descriptivo bajo el enfoque cuantitativo. Para la construcción de la herramienta web Saber Expo BI 5 de limpieza, transformación y visualización de datos se utilizó la metodología ágil de desarrollo Scrum el cual mediante los Sprints permite realizar una investigación ágil y eficiente. Para la herramienta de inteligencia de negocios se propuso 5 Sprints que se profundizan posteriormente.

SPRINT	DETALLES
SPRINT 1. Planeación.	La apropiación de los conceptos básicos para el desarrollo del proyecto tales como: Inteligencia de Negocios. Mercado de datos. Cuadros de mando. Análisis multidimensional OLAP. Diseño y desarrollo de cuadros de mandos.
SPRINT 2. Diseño herramienta.	Se determinó los diferentes tipos de gráficos más adecuados y el maquetado de la vista en virtud del concepto de cada indicador, además se realizó el completo análisis del comportamiento del sistema.

SPRINT 3. Desarrollo herramienta.	<p>En este Sprint se comenzó a definir los distintos paquetes necesarios para el desarrollo de las funcionalidades del aplicativo iniciando con todo lo referente a la visualización de los indicadores que abarca la aplicación haciendo uso de la librería Charts.js junto con otras utilidades. Lo siguiente fue la construcción del diseño multidimensional asociado al mercado de datos, realizando un diccionario de datos asociados a cada dimensión del diseño multidimensional, que fueron necesarios para el diseño del sistema de extracción, transformación y carga (ETL) para la creación de las bodegas de datos.</p>
SPRINT 4. Realización de pruebas.	<p>En esta fase se definieron y aplicaron diferentes tipos de pruebas para mejorar la calidad de la aplicación.</p>
SPRINT 5. Implementación y documentación.	<p>Se subió el aplicativo a un servidor el cual se encuentra en constante actualización y mejoramiento gracias a la herramienta Git. Por último se realizó la respectiva documentación con respecto a los manuales de usuario y manuales del programador.</p>

ENTORNO DE DESARROLLO

Para trabajar con el proyecto se necesita tener instalados los siguientes programas y dependencias:

- PostgreSQL
- Python
- Django
- Pycharm
- Visual Studio Code

A continuación, se indica como instalar y configurar correctamente cada uno de ellos.

PostgreSQL

PostgreSQL es el motor de base de datos de la actualidad. Lo que lo hace el Sistema Gestor de Bases de Datos de Código libre más potente y robusto del mercado.

Ventajas PostgreSQL

- **Instalación ilimitada y gratuita:** Podemos instalarlo en todos los equipos que se desee, independientemente de la plataforma y la arquitectura que usemos, PostgreSQL está disponible para los diferentes SO, Unix, Linux y Windows, en 32 y 64 bits, lo que lo hace un sistema multiplataforma y de gran escalado.
- **Gran escalabilidad:** Nos permite configurar PostgreSQL en cada equipo según el hardware; Por esto, es capaz de ajustarse al número de CPU y a la cantidad de memoria disponible de forma óptima. Con ello logramos una mayor cantidad de peticiones simultáneas a la base de datos de forma correcta.
- **Estabilidad y confiabilidad:** Tiene más de 20 años de desarrollo activo y en constante mejora. No se han presentado nunca caídas de la base de datos que estén documentadas gracias a su capacidad de establecer un entorno de Alta disponibilidad

en donde podemos hacer tareas de mantenimiento o recuperación sin bloquear completamente el sistema.

- **Potencia y Robustez:** PostgreSQL cumple en su totalidad con la característica ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad en español). Por ello permite que las transacciones no interfieran unas con otras. Con ello se garantiza la información de las Bases de Datos y que los datos perduren en el sistema.
- **Extensibilidad:** tenemos a nuestra disponibilidad una gran variedad de extensiones distribuidas por el grupo de desarrolladores de PostgreSQL. Estas extensiones pueden ser lenguajes de programación, tales como, Perl, Java, Python, C++ y muchos más.

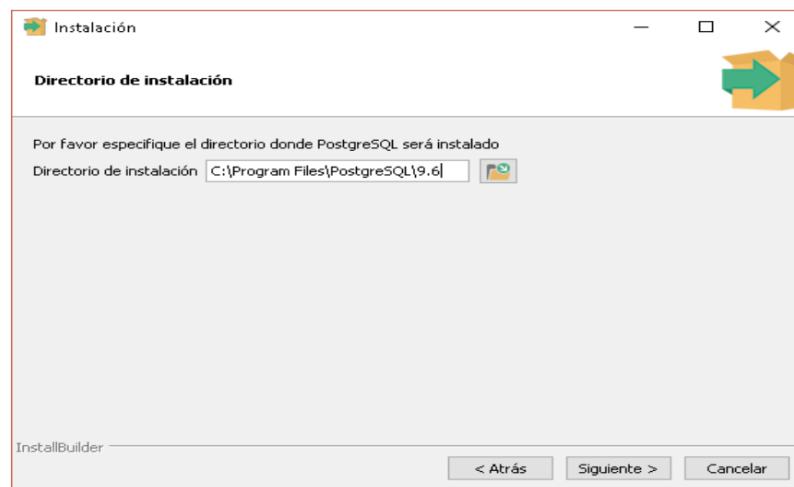
Desventajas de PostgreSQL

- **Base de datos pequeñas:** Es relativamente lento en inserciones y actualizaciones en bases de datos pequeñas, PostgreSQL está diseñado para ambientes de alto volumen, por ello la velocidad de respuesta pueda parecer lenta en comparación con bases de datos de pequeño tamaño.
- **Soporte oficial:** No cuenta con un soporte en línea o telefónico, por el contrario, cuenta con foros oficiales donde los usuarios pueden exponer sus dudas que responden, es de resaltar que la comunidad PostgreSQL es una de las más activas del mercado.
- **Sintaxis:** La sintaxis de algunos de sus comando o sentencias puede llegar a no ser intuitiva si no tienes un nivel medio de conocimientos en lenguaje SQL.

Instalación de PostgreSQL

- Para instalar PostgreSQL nos dirigimos a nuestro navegador y accedemos a la siguiente URL: <https://www.enterprisedb.com/downloads/postgres-postgresql-downloads> nos direccionará a otra página donde podremos descargarnos la versión correcta tanto para nuestras necesidades como para el tipo de sistema operativo que tengamos instalado.

- En nuestro caso es la versión 13.1 para Windows de 64 bits. Una vez descargado el archivo de instalación lo ejecutamos y si tenemos activado el control de cuentas de usuario nos mostrará una advertencia con el texto "¿Desea permitir que este programa realice cambios en el equipo?", pulsaremos "Sí" para continuar con la instalación de PostgreSQL; Se iniciará el asistente para instalar PostgreSQL, pulsaremos "Siguiente".
- Indicaremos los complementos que queremos instalar y la carpeta de instalación de PostgreSQL, donde se guardarán los ejecutables, librerías y ficheros de configuración de PostgreSQL.



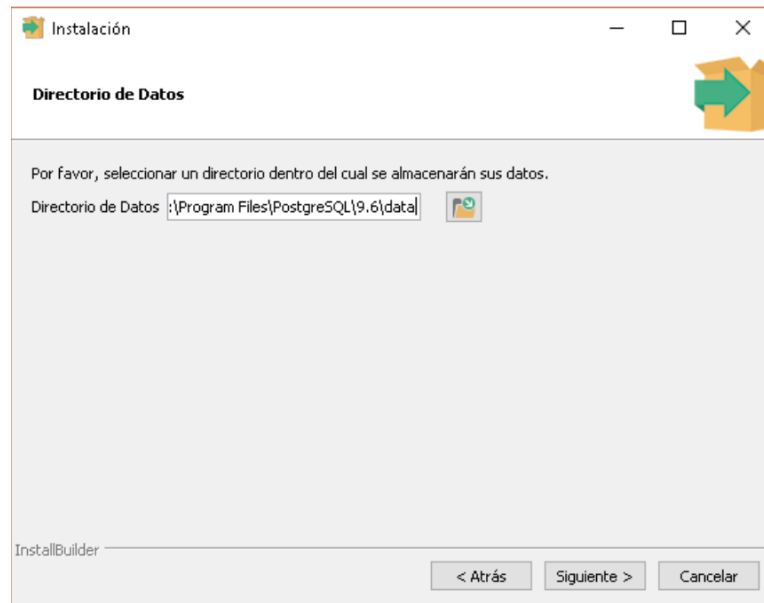
enterpriseDB.com/downloads/postgres-postgresql-downloads

EDB ¿Por qué EDB? Productos Servicios Apoyo Recursos Planes Contacto Regístrate Descargas

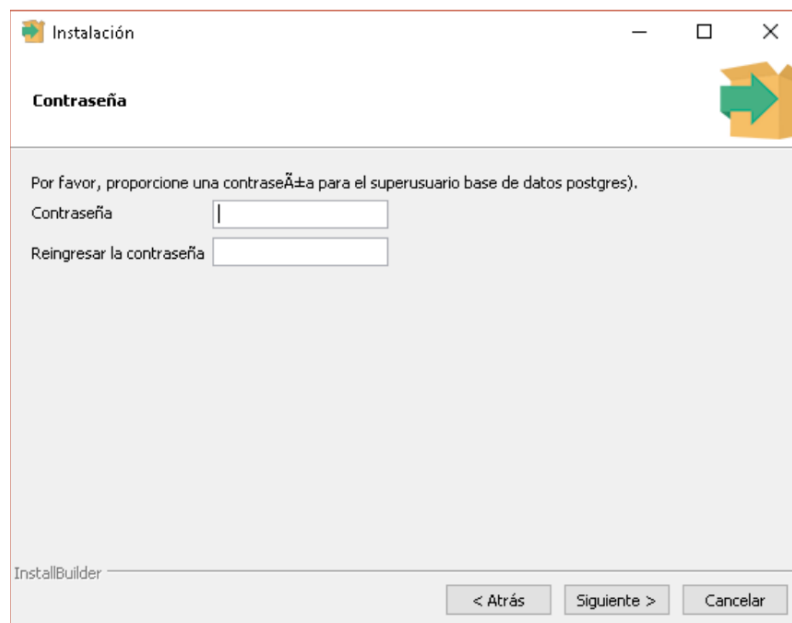
Descarga de la base de datos PostgreSQL

Versión	Linux x86-64	Linux x86-32	Mac OS X	Windows x86-64	Windows x86-32
13.1	N / A	N / A	Descargar	Descargar	N / A
12.5	N / A	N / A	Descargar	Descargar	N / A
11.10	N / A	N / A	Descargar	Descargar	N / A
10.15	Descargar	Descargar	Descargar	Descargar	Descargar
9.6.20	Descargar	Descargar	Descargar	Descargar	Descargar
9.5.24	Descargar	Descargar	Descargar	Descargar	Descargar
9.4.26 (no compatible)	Descargar	Descargar	Descargar	Descargar	Descargar
9.3.25 (no compatible)	Descargar	Descargar	Descargar	Descargar	Descargar

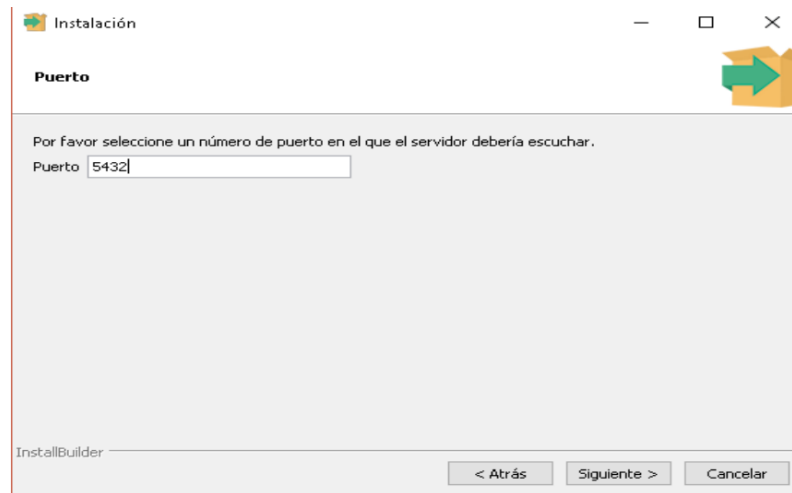
- Indicaremos también la carpeta donde se guardarán los datos por defecto de PostgreSQL.



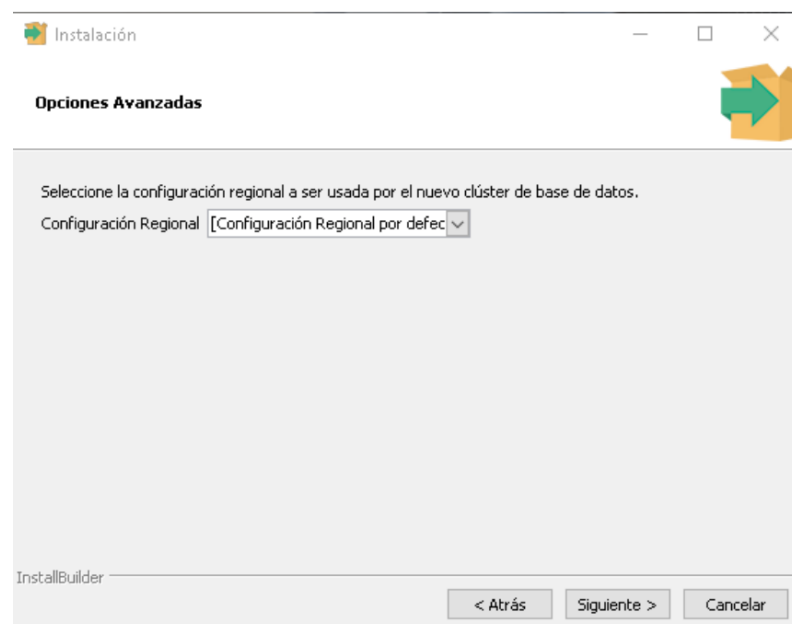
- Introduciremos la contraseña para el súper usuario "postgres" que será con el que iniciemos sesión para administrar la base de datos, si tenemos clave de cuenta en Windows introduciremos nuestra clave.



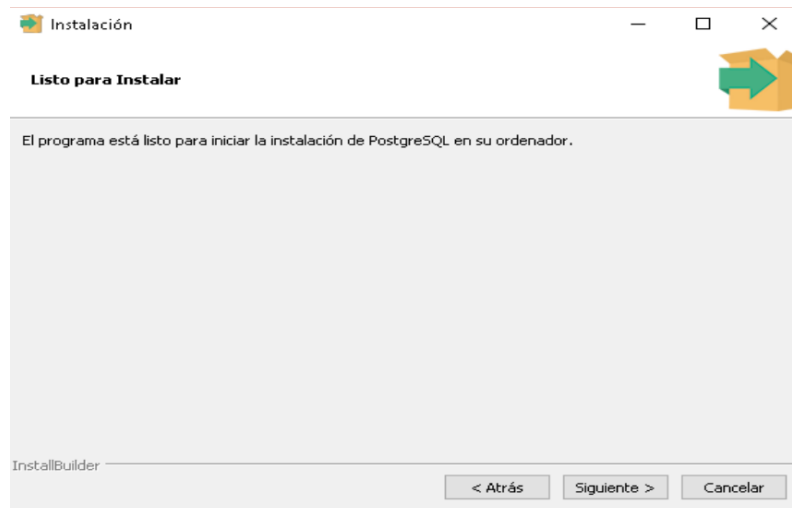
- Introduciremos el puerto de escucha para la conexión con el servidor PostgreSQL, por defecto el 5432.



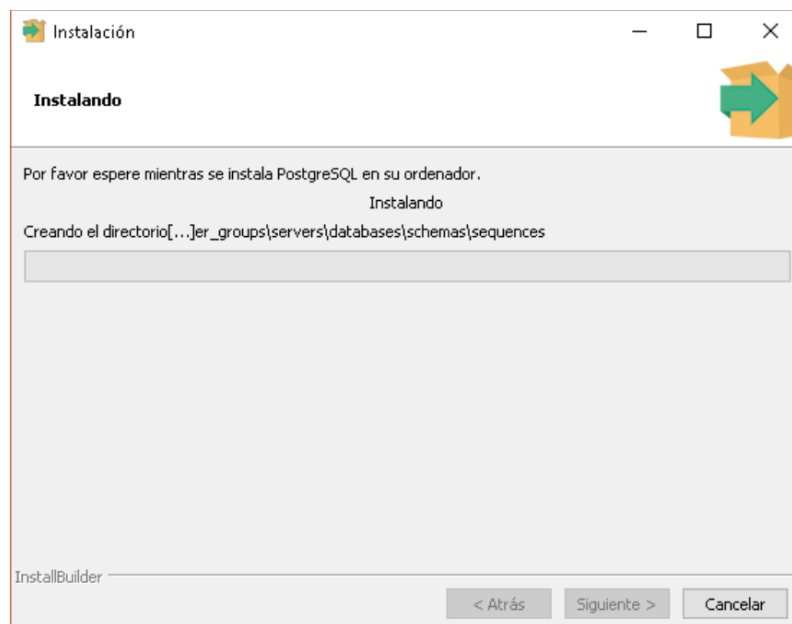
- Seleccionaremos la configuración regional.



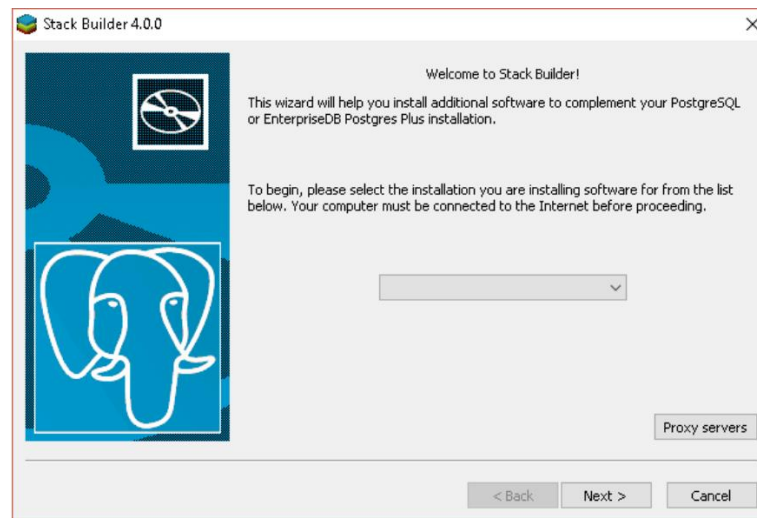
- Pulsaremos "Siguiete" para iniciar la instalación definitiva.



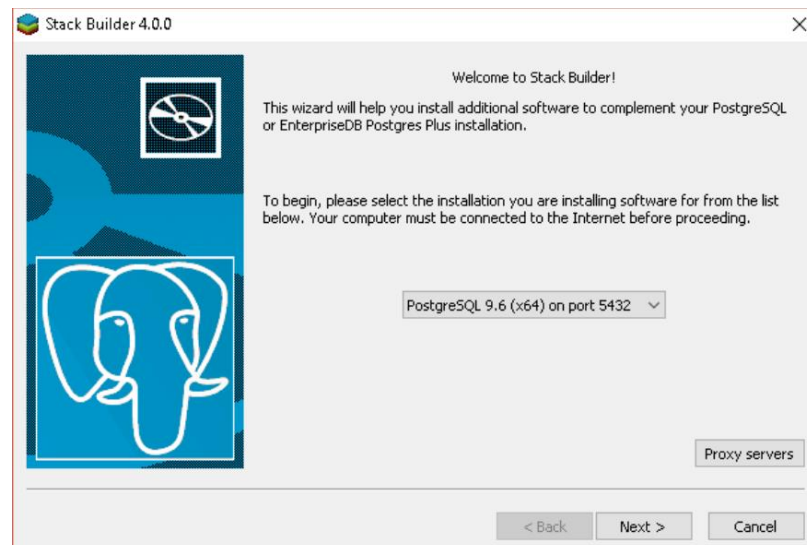
- Se iniciará el asistente para instalar el motor de base de datos PostgreSQL, que creará las carpetas oportunas, copiará los ficheros necesarios y creará el servicio Windows para iniciar de forma automática el motor de base de datos.



- Una vez finalizada la instalación el asistente nos dará la posibilidad de ejecutar Stack Builder, aplicación que nos permitirá instalar otros componentes y herramientas para PostgreSQL.



- Si hemos marcado la opción de Stack Builder, se iniciará, seleccionaremos "PostgreSQL on port 5432". Pulsaremos "Finish" (en nuestro caso cancelaremos Stack Builder pues no instalaremos más componentes).



Python

Es un lenguaje de programación versátil multiplataforma y multiparadigma que se destaca por su código legible y limpio. Una de las razones de su éxito es que cuenta con una licencia de código abierto que permite su utilización en cualquier escenario.

Python es ideal para trabajar con grandes volúmenes de datos ya que, al ser multiplataforma, favorece su extracción y procesamiento, por eso lo eligen las empresas de Big Data. A nivel científico, tiene una gran biblioteca de recursos con especial énfasis en las matemáticas para aspirantes a programadores en áreas especializadas. Y si esto fuera poco, es posible crear videojuegos, aunque no es tan eficiente como Java o C#.

Ventajas de Python

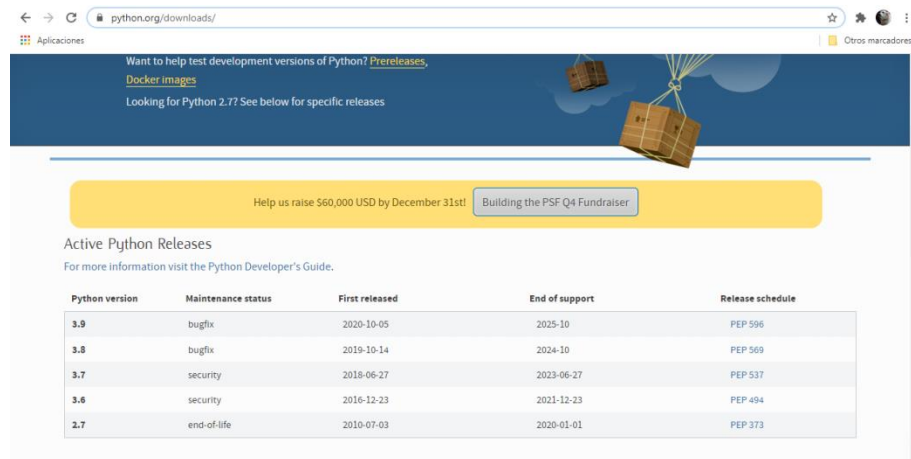
- **Simplificado y rápido:** Este lenguaje simplifica mucho la programación, es un gran lenguaje para scripting.
- **Elegante y flexible:** El lenguaje ofrece muchas facilidades al programador al ser fácilmente legible e interpretable.
- **Programación sana y productiva:** Es sencillo de aprender, con una curva de aprendizaje moderada. Es muy fácil comenzar a programar y fomenta la productividad.
- **Ordenado y limpio:** es muy legible y sus módulos están bien organizados.
- **Portable:** Es un lenguaje muy portable. Podemos usarlo en prácticamente cualquier sistema de la actualidad.
- **Comunidad:** Cuenta con un gran número de usuarios. Su comunidad participa activamente en el desarrollo del lenguaje.

Desventajas de Python

- **Hosting:** La mayoría de los servidores no tienen soporte a Python, y si lo soportan, la configuración es un poco difícil.
- **Librerías incluidas:** Algunas librerías que trae por defecto no son del gusto de amplio de la comunidad, y optan a usar librerías de terceros.

Instalación de Python

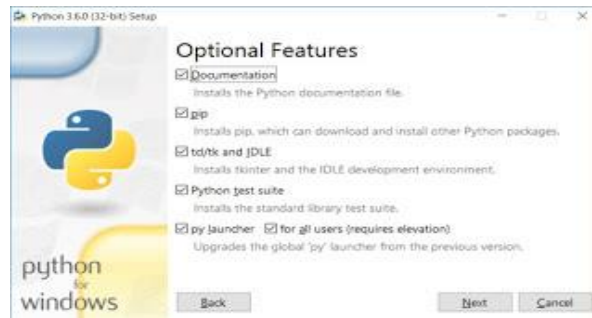
- Para la instalación accedemos desde nuestro navegador al sitio oficial de Python desde la siguiente URL <https://www.python.org/downloads/> y descargamos el instalador ejecutable de Python según los requerimientos de nuestro computador.



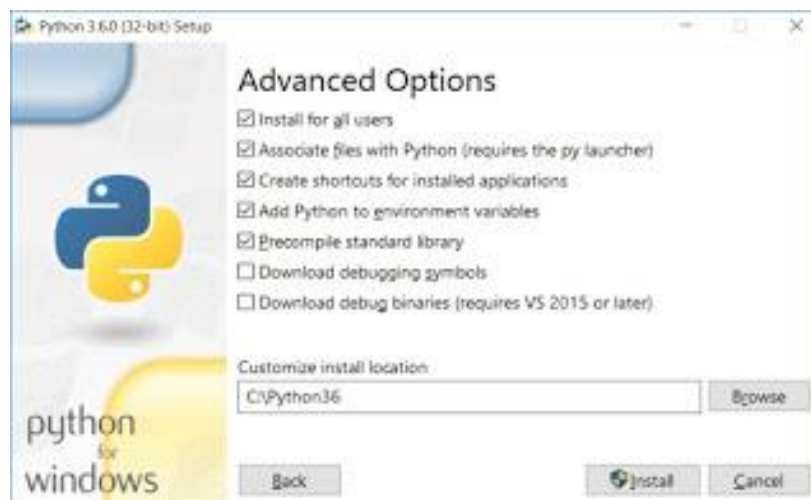
- Iniciar la instalación ejecutando el archivo descargado Python.exe con doble clic. Si es necesario, confirmar la ejecución en la ventana de advertencia, una vez iniciado el instalador, en la ventana Install Python activar las casillas de las opciones: Install launcher for all users (recommended) y Add Python 3.6 to PATH. Después, continuar seleccionando la opción Customize installation. Choose Location and features.



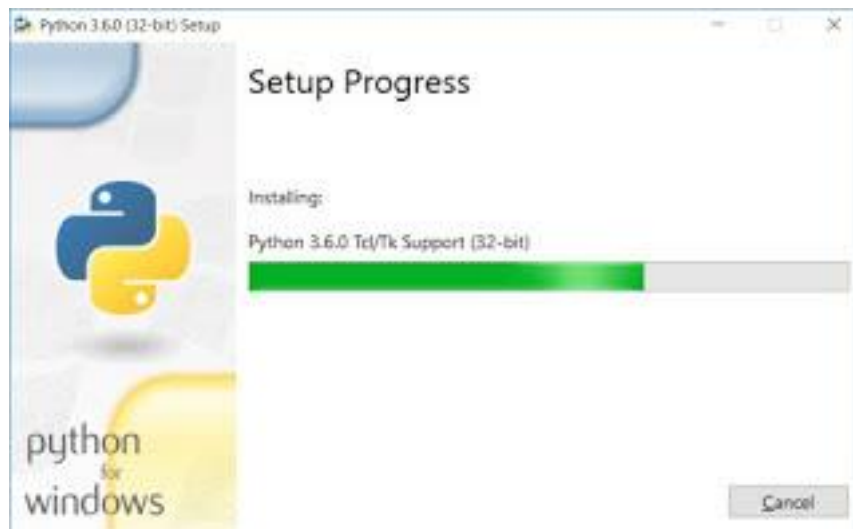
- En la ventana Optional features verificar que están activas todas las opciones y hacer clic en el botón Next.



- En la ventana Advanced Options verificar que están activas las opciones de la imagen, escribir la ruta del directorio de instalación "C:\Python36" (o aceptar la ruta por defecto) y comenzar la instalación haciendo clic en el botón Install.



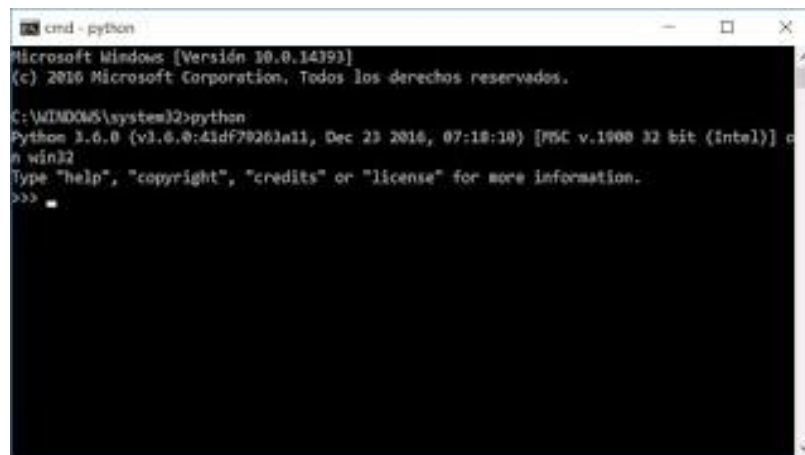
- A continuación, después de unos segundos de espera, comenzará el proceso de instalación de Python.



- En la ventana Setup was successful, una vez que ha concluido el proceso de instalación hacer clic en el botón Close. Desde esta ventana es posible acceder a un tutorial online de Python, a la documentación oficial del lenguaje y a información con las novedades de la presente versión.



- Para probar Python acceder a la línea de comandos ejecutando el programa cmd.exe o Símbolo del sistema. Después, escribir "python" y aceptar para iniciar el entorno interactivo de Python donde se puede verificar la versión de Python instalada y es posible introducir comandos Python. Para finalizar, introducir quit() o presionar Ctrl+D.



Django

Django es un framework web de alto nivel que permite el desarrollo rápido de sitios web seguros y mantenibles. Desarrollado por programadores experimentados, Django se encarga

de gran parte de las complicaciones del desarrollo web, por lo que puedes concentrarte en escribir tu aplicación sin necesidad de reinventar la rueda. Es gratuito y de código abierto, tiene una comunidad próspera y activa, una gran documentación y muchas opciones de soporte gratuito y de pago.

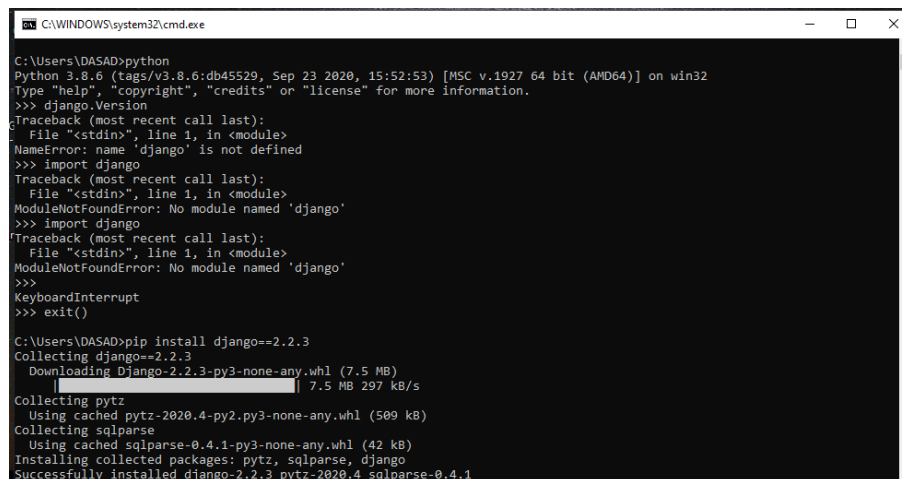
Características:

- **Completo:** Django sigue la filosofía "Baterías incluidas" y provee casi todo lo que los desarrolladores quisieran que tenga "de fábrica". Porque todo lo que necesitas es parte de un único "producto", todo funciona a la perfección, sigue principios de diseño consistentes y tiene una amplia y actualizada documentación.
- **Versátil:** Django puede ser para construir casi cualquier tipo de sitio web, desde sistemas manejadores de contenidos y wikis, hasta redes sociales y sitios de noticias. Puede funcionar con cualquier framework en el lado del cliente.
- **Seguro:** Django ayuda a los desarrolladores evitar varios errores comunes de seguridad al proveer un framework que ha sido diseñado para "hacer lo correcto" para proteger el sitio web automáticamente. Por ejemplo, Django, proporciona una manera segura de administrar cuentas de usuario y contraseñas, evitando así errores comunes como colocar informaciones de sesión en cookies donde es vulnerable
- **Escalable:** Tiene en cuenta una clara separación entre las diferentes partes significa que puede escalar para aumentar el tráfico al agregar hardware en cualquier nivel, servidores de cache, servidores de bases de datos o servidores de aplicación.
- **Mantenible:** El código de Django está escrito usando principios y patrones de diseño para fomentar la creación de código mantenible y reutilizable. En particular, utiliza el principio No te repitas "Don't Repeat Yourself" (DRY) para que no exista una duplicación innecesaria, reduciendo la cantidad de código.

- **Portable:** Django está escrito en Python, el cual se ejecuta en muchas plataformas. Lo que significa que no está sujeto a ninguna plataforma en particular, y puede ejecutar sus aplicaciones en muchas distribuciones de Linux, Windows y Mac OS X.

Instalación de Django

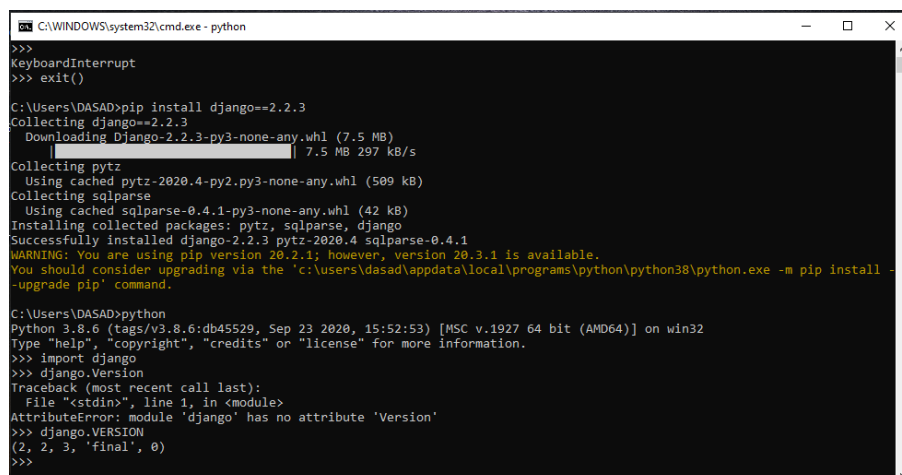
Para instalar Django, se debió instalar Python previamente, pues este es un framework que está escrito en Python y basta con abrir el intérprete de comandos de Windows y escribir el comando “pip install django==2.2.3” y se comenzaran a descargar los repositorios de Django en nuestro ordenador.



```
C:\WINDOWS\system32\cmd.exe
C:\Users\DASAD>python
Python 3.8.6 (tags/v3.8.6:db45529, Sep 23 2020, 15:52:53) [MSC v.1927 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> django.Version
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
NameError: name 'django' is not defined
>>> import django
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ModuleNotFoundError: No module named 'django'
>>> import django
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ModuleNotFoundError: No module named 'django'
>>>
KeyboardInterrupt
>>> exit()

C:\Users\DASAD>pip install django==2.2.3
Collecting django==2.2.3
  Downloading Django-2.2.3-py3-none-any.whl (7.5 MB)
    | 7.5 MB 297 kB/s
Collecting pytz
  Using cached pytz-2020.4-py2.py3-none-any.whl (509 kB)
Collecting sqlparse
  Using cached sqlparse-0.4.1-py3-none-any.whl (42 kB)
Installing collected packages: pytz, sqlparse, django
Successfully installed django-2.2.3 pytz-2020.4 sqlparse-0.4.1
```

Para comprobar que Django esté instalado escribimos en el interprete de comandos la palabra “Python” y dentro de Python escribimos “import django” y luego “django.VERSION”, si esta todo correcto podremos visualizar la versión de Django que hemos instalado.



```
C:\WINDOWS\system32\cmd.exe - python
>>>
KeyboardInterrupt
>>> exit()

C:\Users\DASAD>pip install django==2.2.3
Collecting django==2.2.3
  Downloading Django-2.2.3-py3-none-any.whl (7.5 MB)
    | 7.5 MB 297 kB/s
Collecting pytz
  Using cached pytz-2020.4-py2.py3-none-any.whl (509 kB)
Collecting sqlparse
  Using cached sqlparse-0.4.1-py3-none-any.whl (42 kB)
Installing collected packages: pytz, sqlparse, django
Successfully installed django-2.2.3 pytz-2020.4 sqlparse-0.4.1
WARNING: You are using pip version 20.2.1; however, version 20.3.1 is available.
You should consider upgrading via the 'c:\users\dasad\appdata\local\programs\python\python38\python.exe -m pip install --upgrade pip' command.

C:\Users\DASAD>python
Python 3.8.6 (tags/v3.8.6:db45529, Sep 23 2020, 15:52:53) [MSC v.1927 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> import django
>>> django.Version
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
AttributeError: module 'django' has no attribute 'Version'
>>> django.VERSION
(2, 2, 3, 'final', 0)
>>>
```

PyCharm

Es un IDE o entorno de desarrollo integrado multiplataforma utilizado para desarrollar en el lenguaje de programación Python. Proporciona análisis de código, depuración gráfica, integración con VCS / DVCS y soporte para el desarrollo web con Django, entre otras bondades. PyCharm es desarrollado por la empresa JetBrains y debido a la naturaleza de sus licencias tiene dos versiones, la Community que es gratuita y orientada a la educación y al desarrollo puro en Python y la Professional, que incluye más características como el soporte a desarrollo web con varios precios.

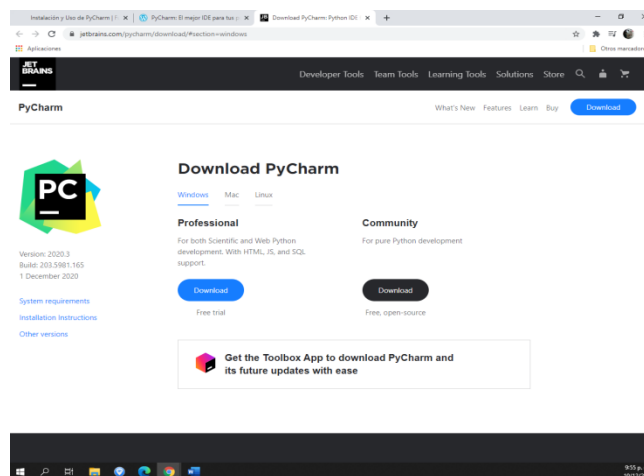
Características:

- Resaltado de sintaxis, sangría automática y formateo de código, finalización de código, formateador de código, resaltado de errores sobre la marcha, análisis del código a medida que se escribe y detector de código duplicado.
- Depurador gráfico en local.
- Refactorizaciones rápidas y seguras válidas para nombres de fichero, funciones, constantes, clases, propiedades, métodos, parámetros y variables locales y globales.
- Busca en todas partes, encontrar usos de cualquier símbolo, ir a clase/archivo/símbolo, ir a declaración, marcadores personalizados.
- Integración con sistemas de control de versiones: Git, Mercurial, CVS, Subversion, GitHub.
- Integración con documentación: reStructuredText y Google.
- Integración con PyQt y PyGTK.

- Gestión de paquetes.
- Integración nativa con Virtualenv/Buildout.
- Consola de Python.
- Soporte para XML, HTML, YAML, JSON, RelaxNG.
- Terminal local.

Instalación de PyCharm

- La versión gratuita Pycharm Community se puede descargar de la siguiente dirección:
<https://www.jetbrains.com/pycharm/download/>

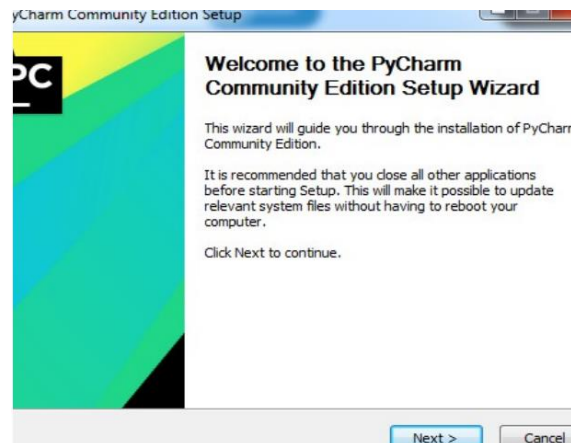


- Para poder usar PyCharm es necesario tener una licencia, podemos elegir la versión community en la cual debemos ingresar una licencia estudiantil que podremos gestionar en la siguiente página, el único requisito es tener una cuenta de correo institucional; la licencia se puede obtener de la siguiente URL:
<https://www.jetbrains.com/es-es/commnity/education/?fromMenu#students>

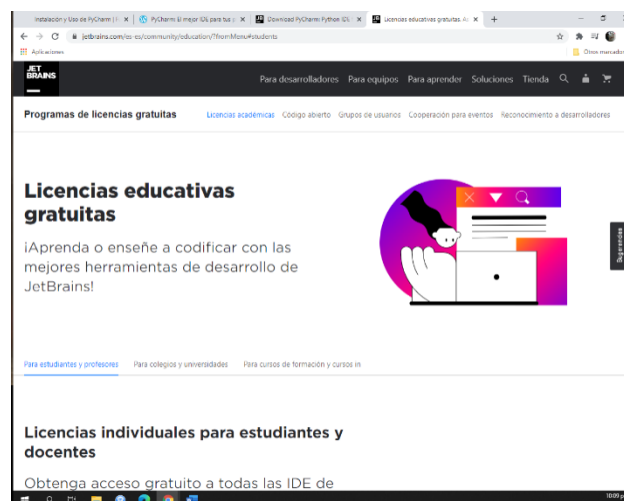
- Ejecuta como administrador el archivo que se acaba de descargar.

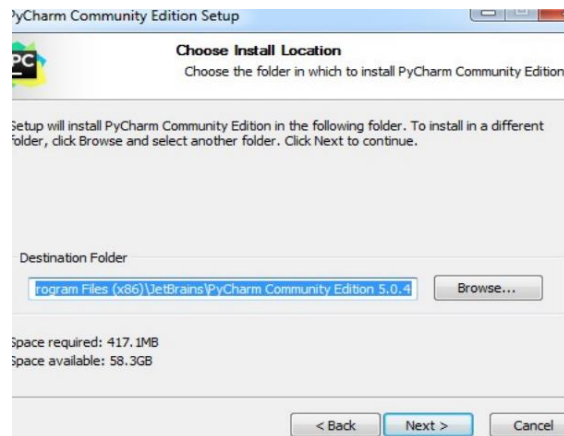


- Pulsar siguiente.

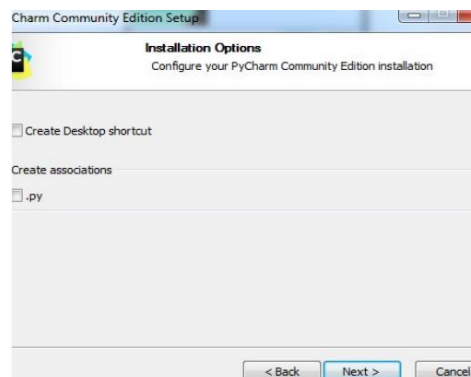


- Elegir la ruta donde se desee instalar y pulsar siguiente.

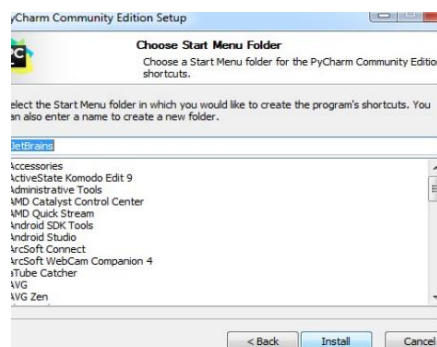




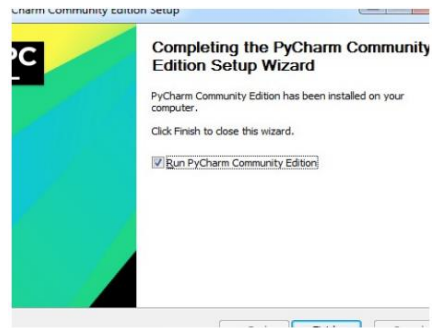
- Pulsar siguiente.



- Seleccionamos la carpeta jetbrains y pulsamos instalar.



- Seleccionamos la opción de correr el programa y pulsamos finalizar.

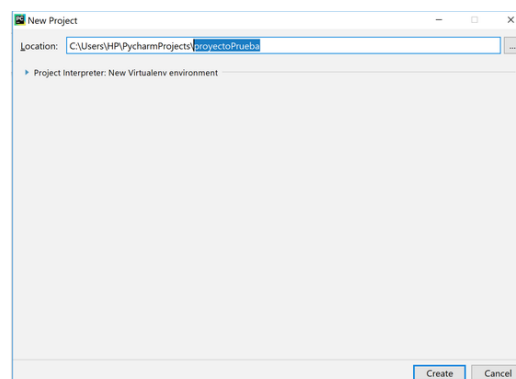


Crear nuevo proyecto PyCharm con Python

- Después de haber instalado PyCharm procederemos a crear un nuevo proyecto.

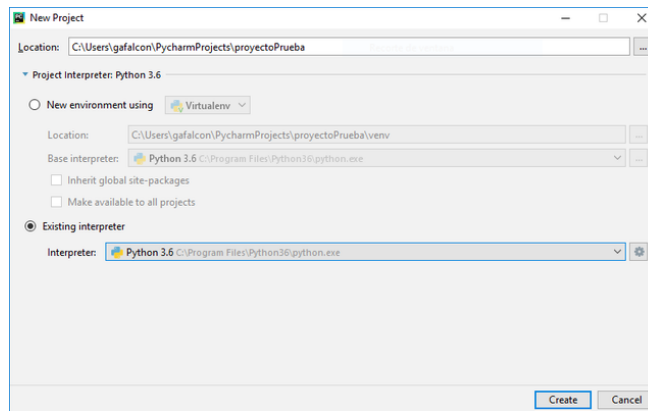


- En la pantalla siguiente, escribir un nuevo nombre a nuestro nuevo proyecto. en este caso, hemos escrito “proyectoPrueba”



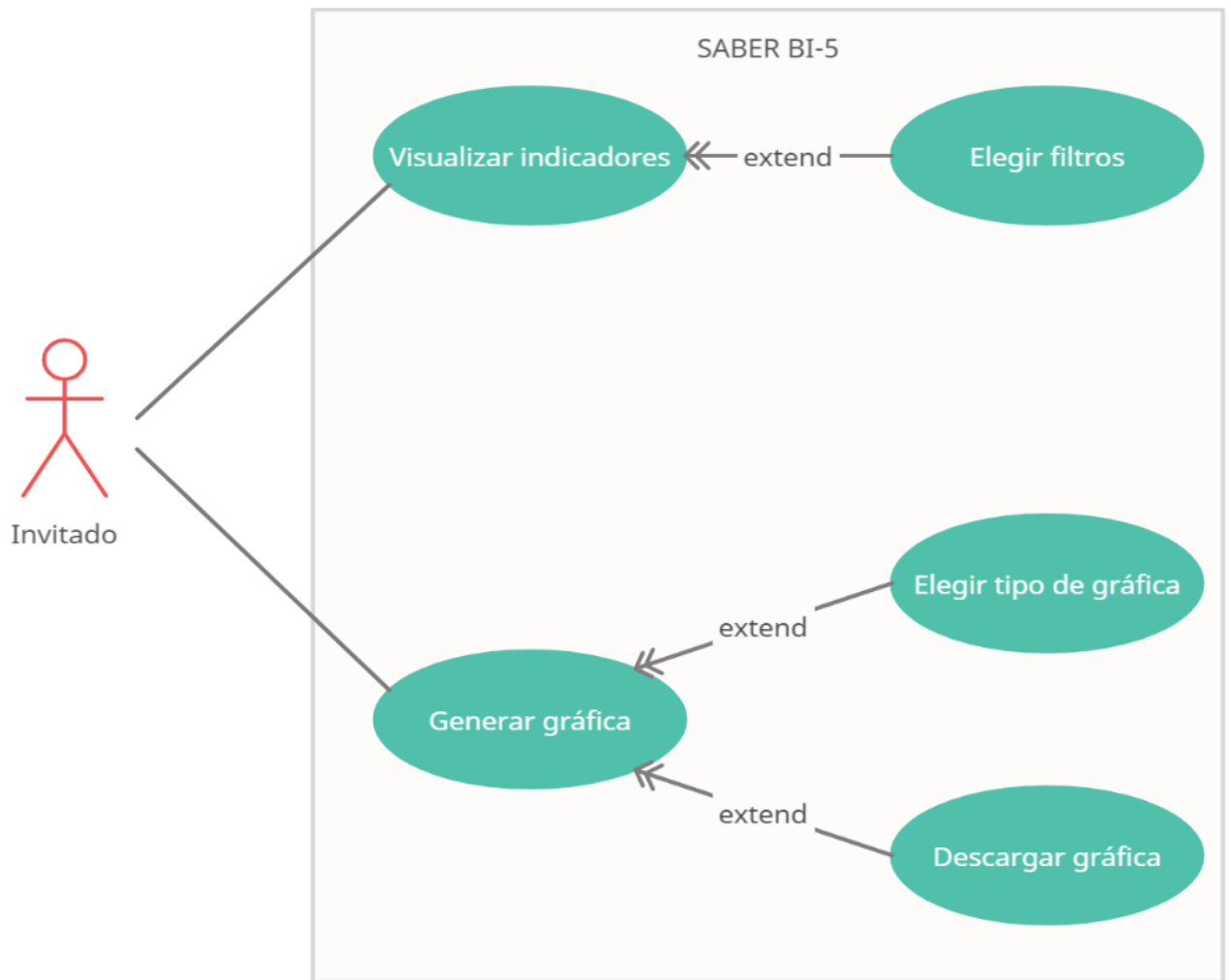
- En la opción Project Interpreter debemos seleccionar el intérprete de Python a utilizar en el proyecto. Pycharm nos presenta dos opciones:

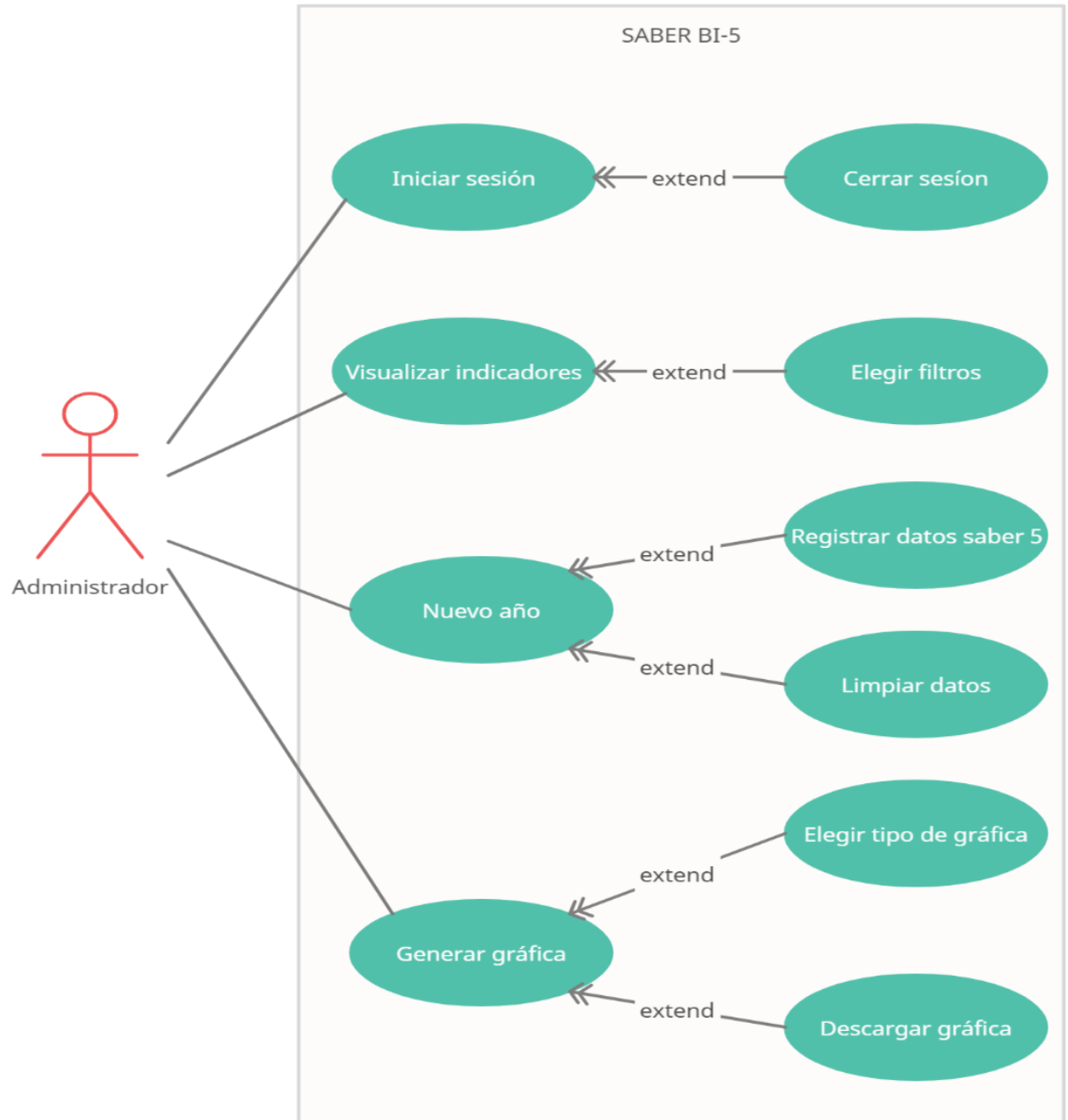
- ✓ Nuevo ambiente (New Environment): Si seleccionamos esta opción, pycharm instalará un intérprete de python exclusivo para el proyecto a crear. De esta manera podremos tener diferentes intérpretes de python para cada proyecto.
 - ✓ Intérprete existente (Recomendado): Pycharm nos permitirá escoger el intérprete de python que tengamos instalado en nuestra computadora. Se recomienda utilizar esta opción que utilizará el intérprete de Python que instalamos previamente.
- Una vez seleccionado el intérprete, hacemos click en “Crear” y se creará un nuevo proyecto.



DISEÑO FUNCIONAL

Las características funcionales del sistema se ilustran por medio de diagramas de Casos de Uso, los cuales se concentran en dos roles que interactúan con el sistema, el usuario general y el usuario administrador representados en los siguientes diagramas.





DISEÑO ESTRUCTURAL

Aquí se explica las características estructurales de un sistema general, la arquitectura y distribución de artefactos de software requerida.



Diagrama De Paquetes

En el desarrollo de la aplicación se maneja Python con un Framework el cual es un entorno Web llamado Django y nos permite trabajar con una estructura MTV (model-template-view),

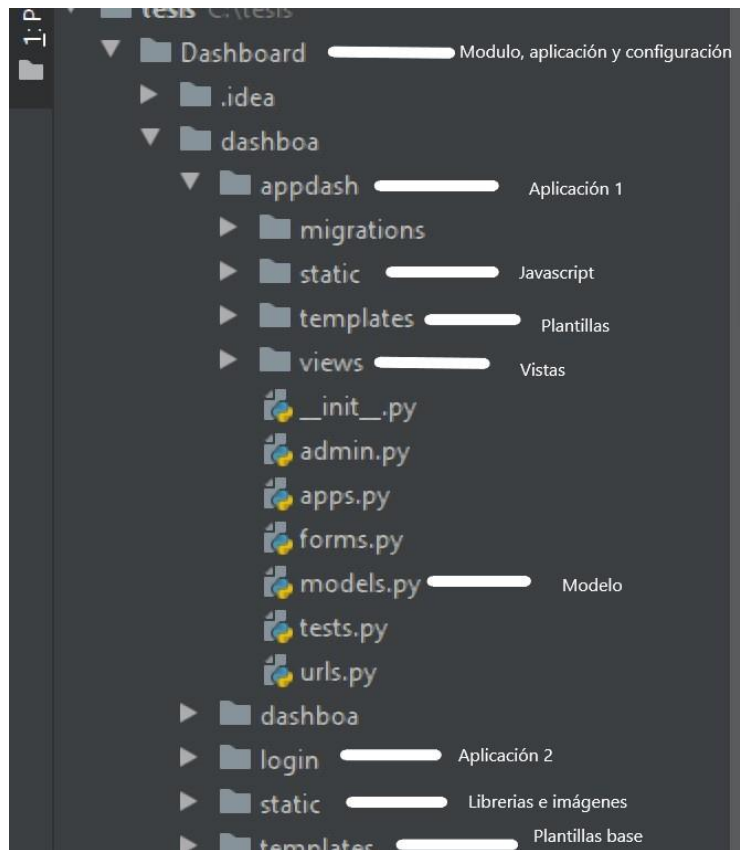
A continuación, se detalla el manejo de dicha estructura:

M = Modelos (Base de datos): Capa de interacción y comunicación que muestra un archivo donde podemos observar las clases que detallan las diferentes dimensiones manejadas en el proyecto.

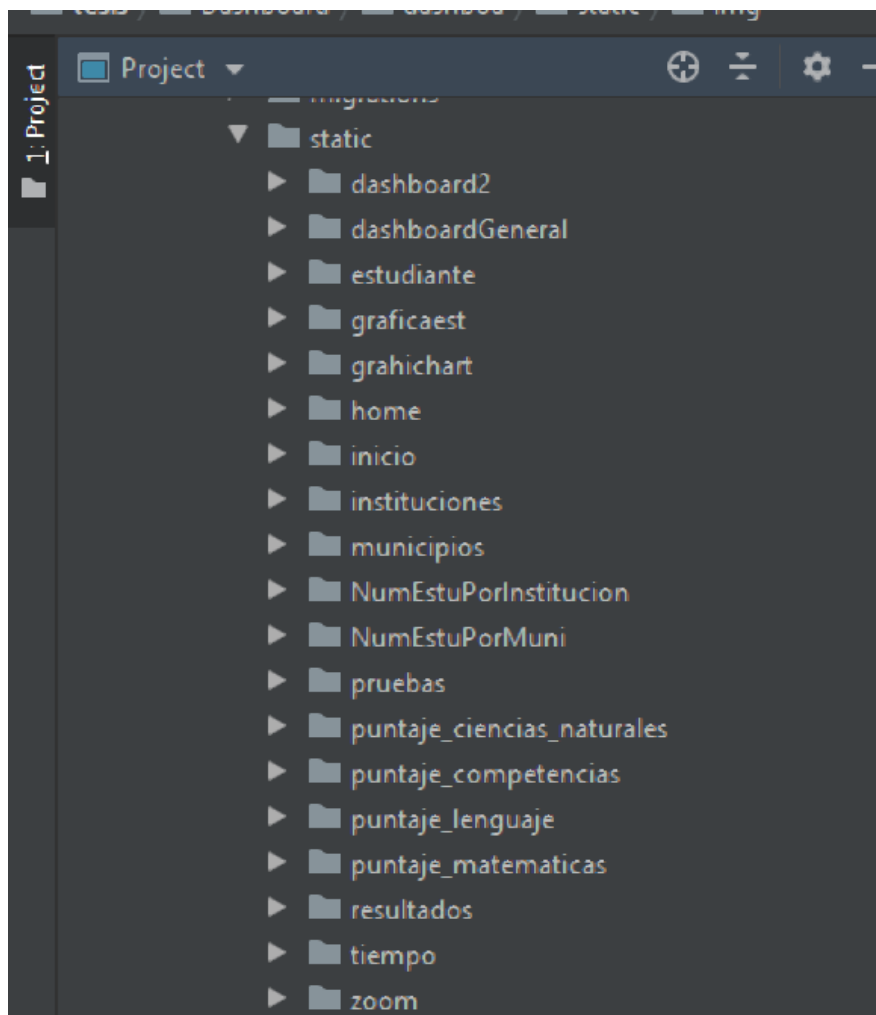
V = Vista (Funciones): permite interactuar con el modelo y es donde encontramos la lógica con la cual realizamos las diferentes operaciones para generar los diferentes resultados.

T = Templates (Plantillas): son los diferentes archivos que se encuentran en la capa de presentación y nos permite mostrar los resultados a los diferentes usuarios. A continuación, se observa las rutas o directorios donde se encuentran cada uno de los diferentes archivos de

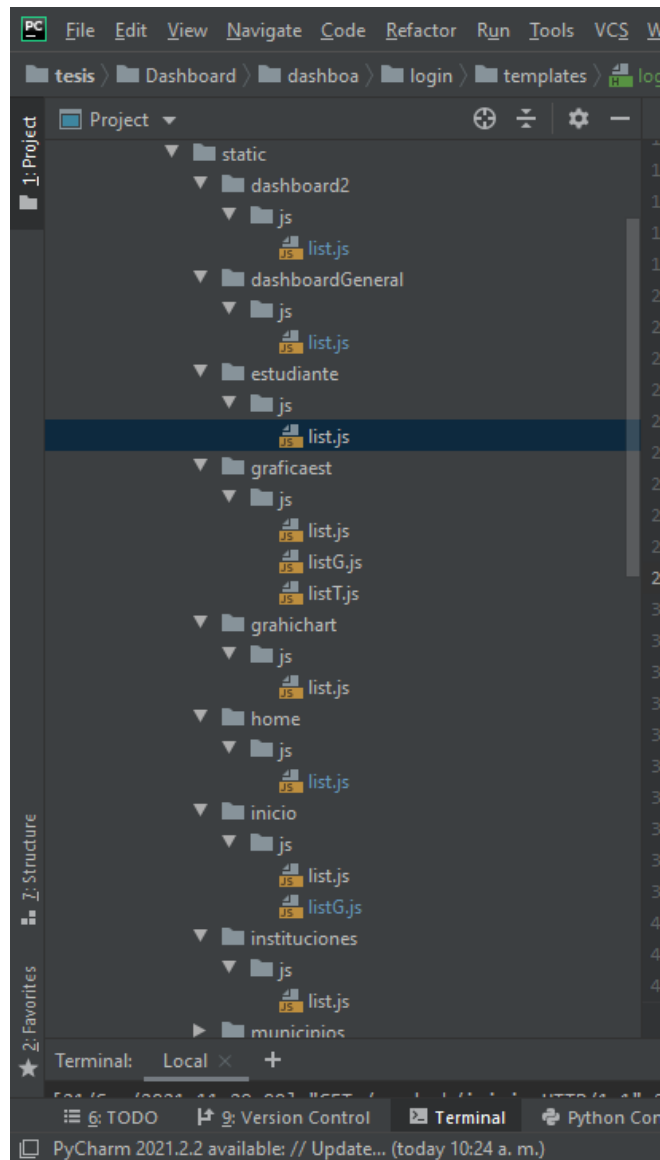
conFiguración para el desarrollo de la aplicación. La imagen nos permite observar de forma general los directorios con sus respectivos nombres en los cuales se encuentran las Plantillas, Vistas, Modelos y Scripts además archivos de conFiguración para su conexión entre los diferentes archivos utilizados en la aplicación.



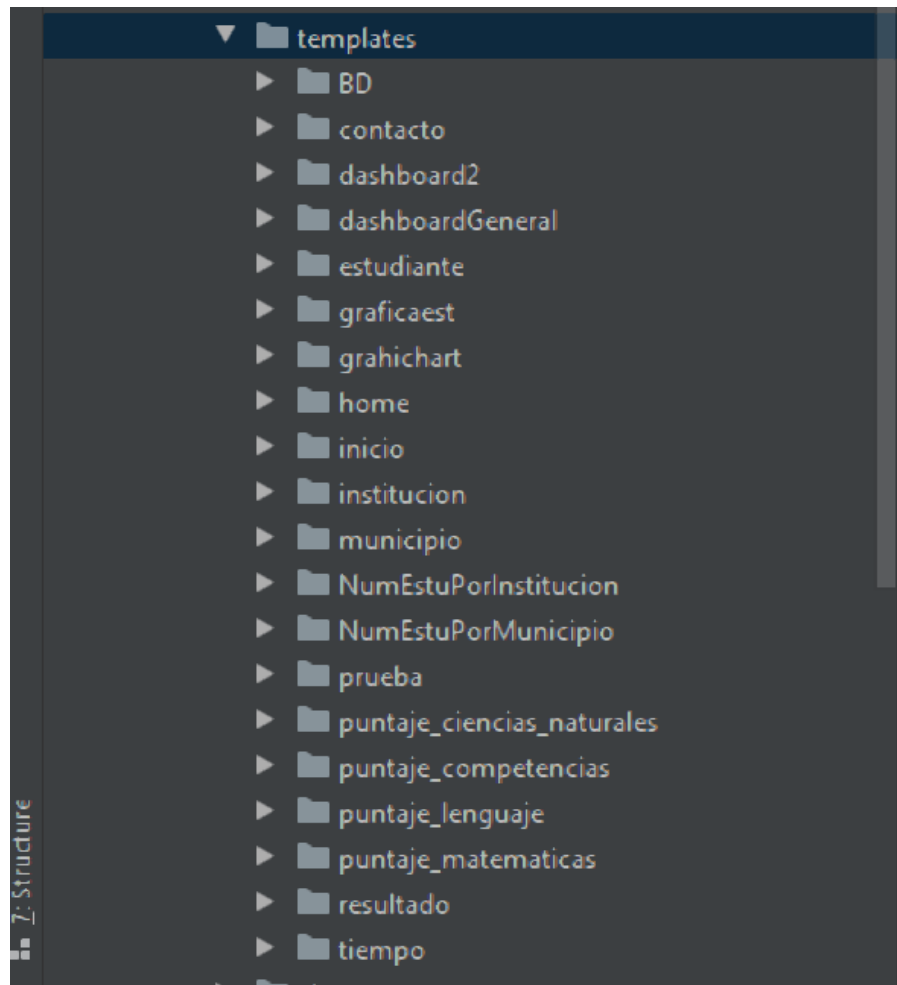
En la siguiente imagen podemos visualizar los diferentes directorios donde se encuentran todos los archivos Scripts para cada plantilla los cuales permite recibir los datos que vienen desde el modelo conectado con la vista y la BD.



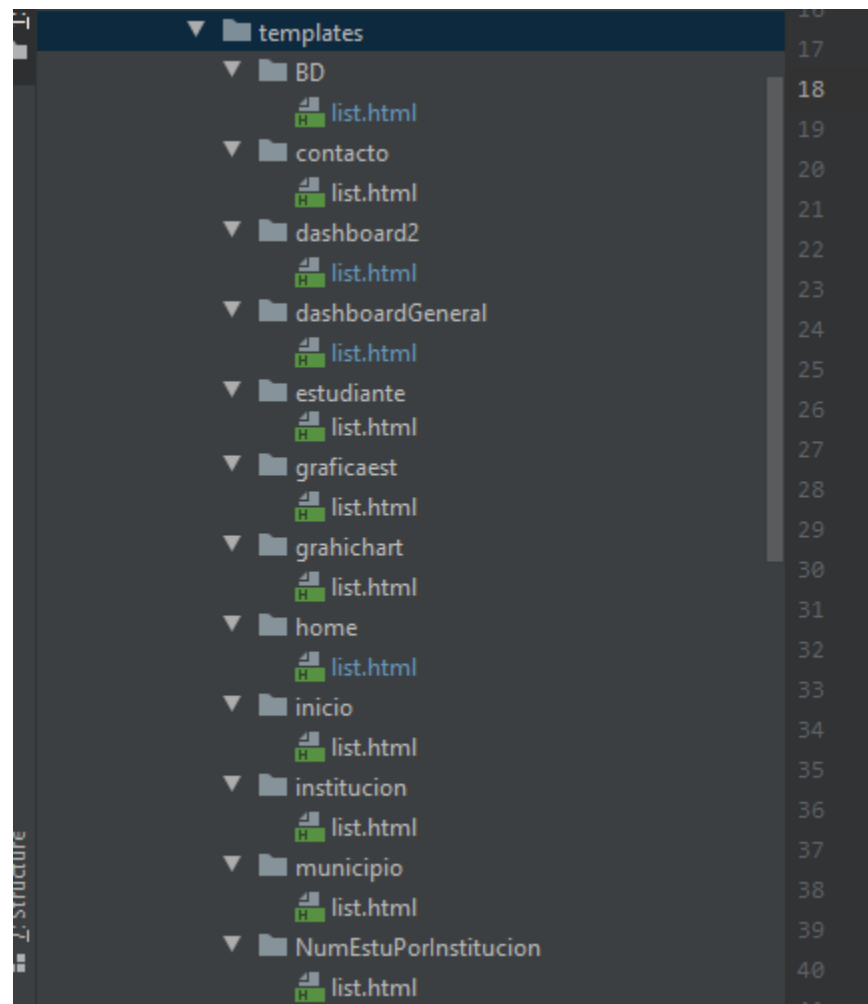
En la siguiente imagen podemos observar desplegados los diferentes directorios que contienen cada uno de los archivos Scripts (JS) que son utilizados con cada plantilla y vista para la visualización de los datos.



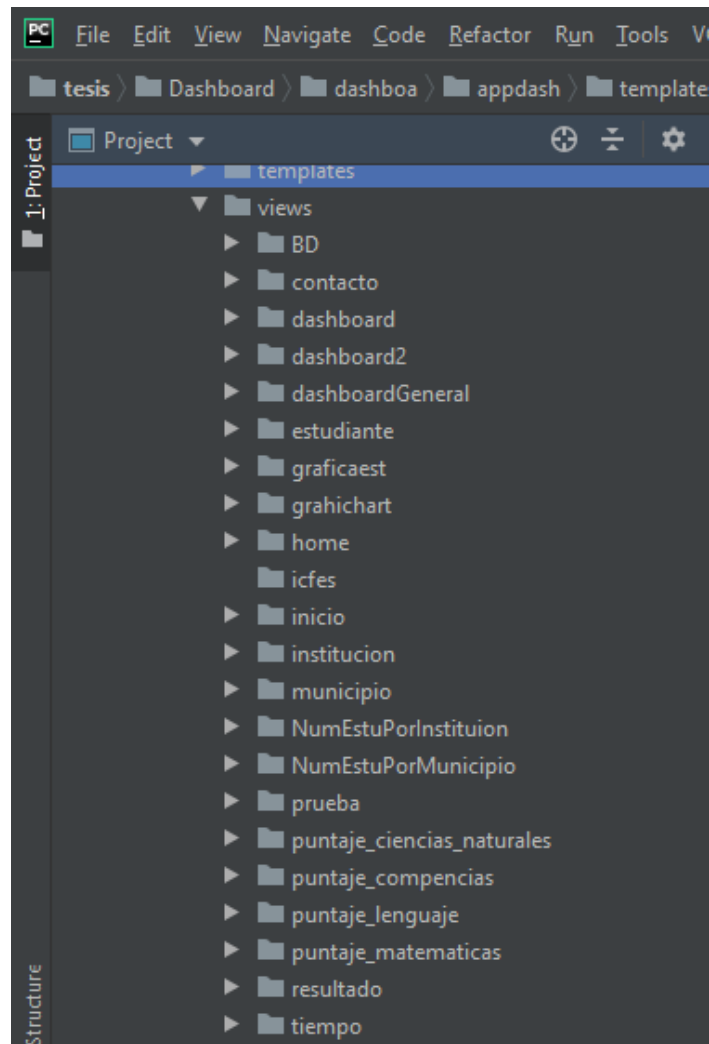
En la siguiente grafica se encuentran las plantillas utilizadas para la visualización de las gráficas y tablas para el usuario.



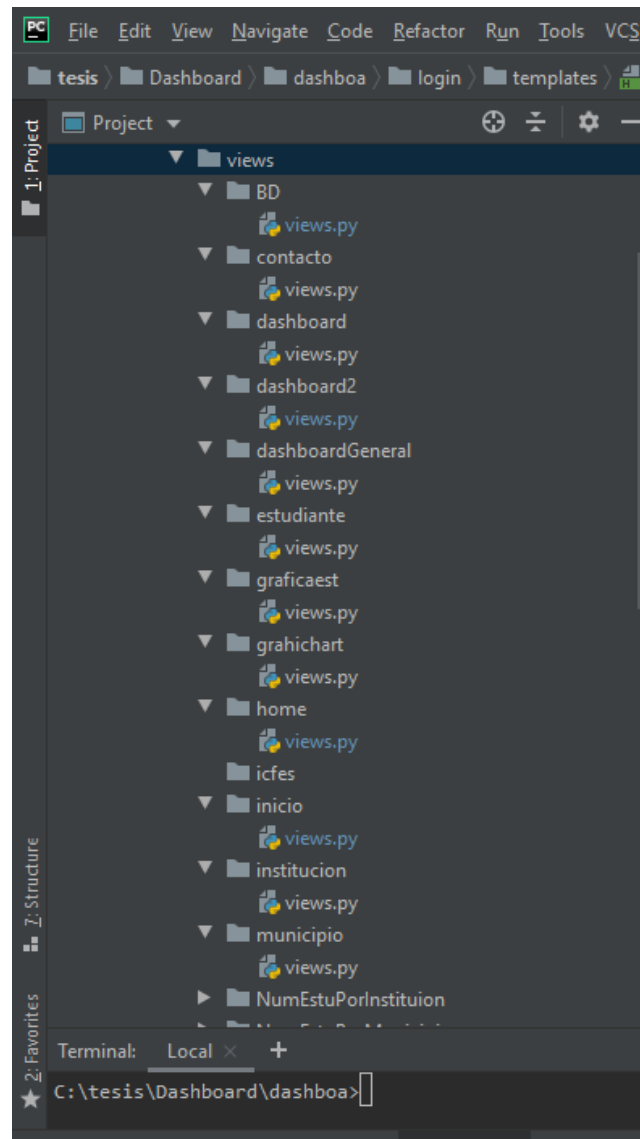
Ahora se visual los diferentes directorios desplegados con sus respectivos archivos con extensión (HTML).



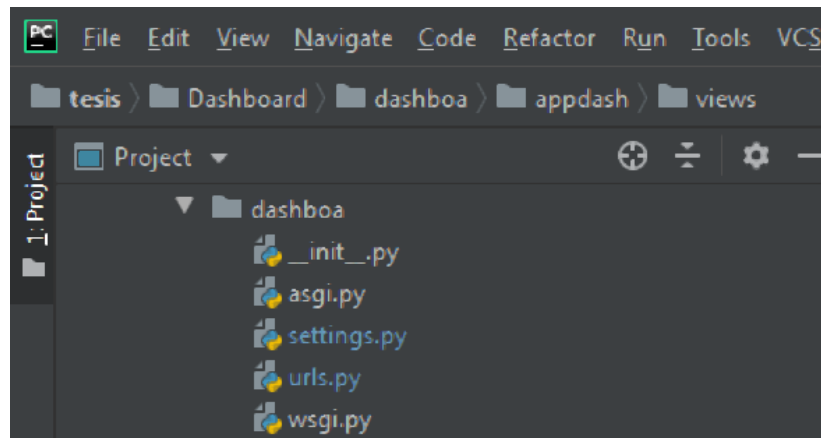
En la siguiente imagen podemos visualizar los directorios que contienen las diferentes vistas utilizadas para el manejo de la lógica en la aplicación.



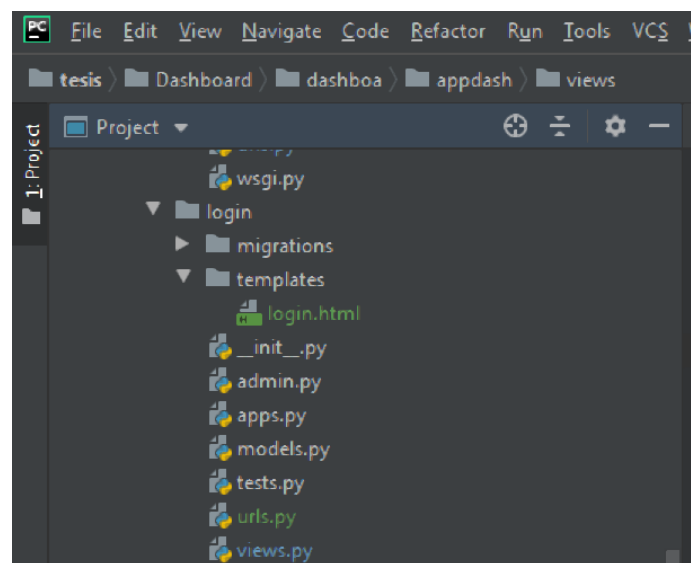
En la siguiente grafica podemos observar desplegados los diferentes directorios donde se encuentras los archivos que contienen las respectivas vistas.



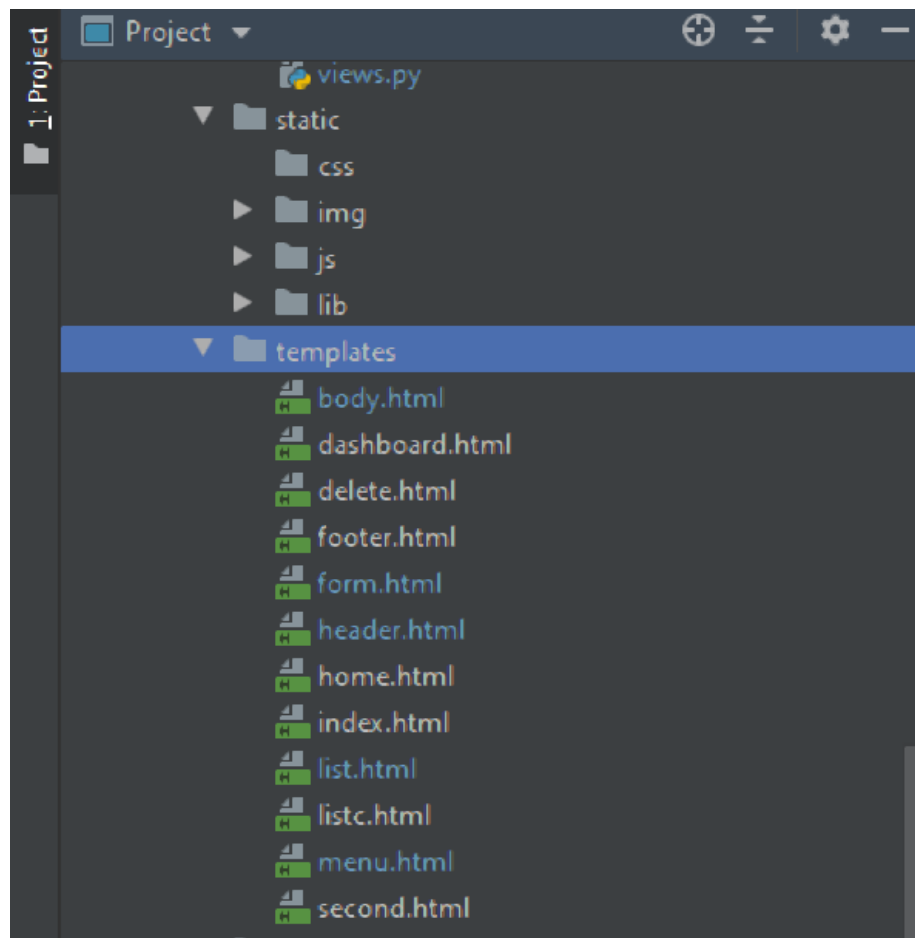
En la siguiente imagen podemos observar el directorio dashboard el cual contiene el archivo urls.py en el cual podemos encontrar las diferentes rutas las cuales permiten conectar con las diferentes vistas para tener acceso a la información de la base de datos según la lógica descrita en cada vista.



En el directorio login podemos encontrar la plantilla en el directorio templates y la vista en el archivo views.py los cuales permiten la configuración del login de usuario para administrador e invitado en la aplicación.



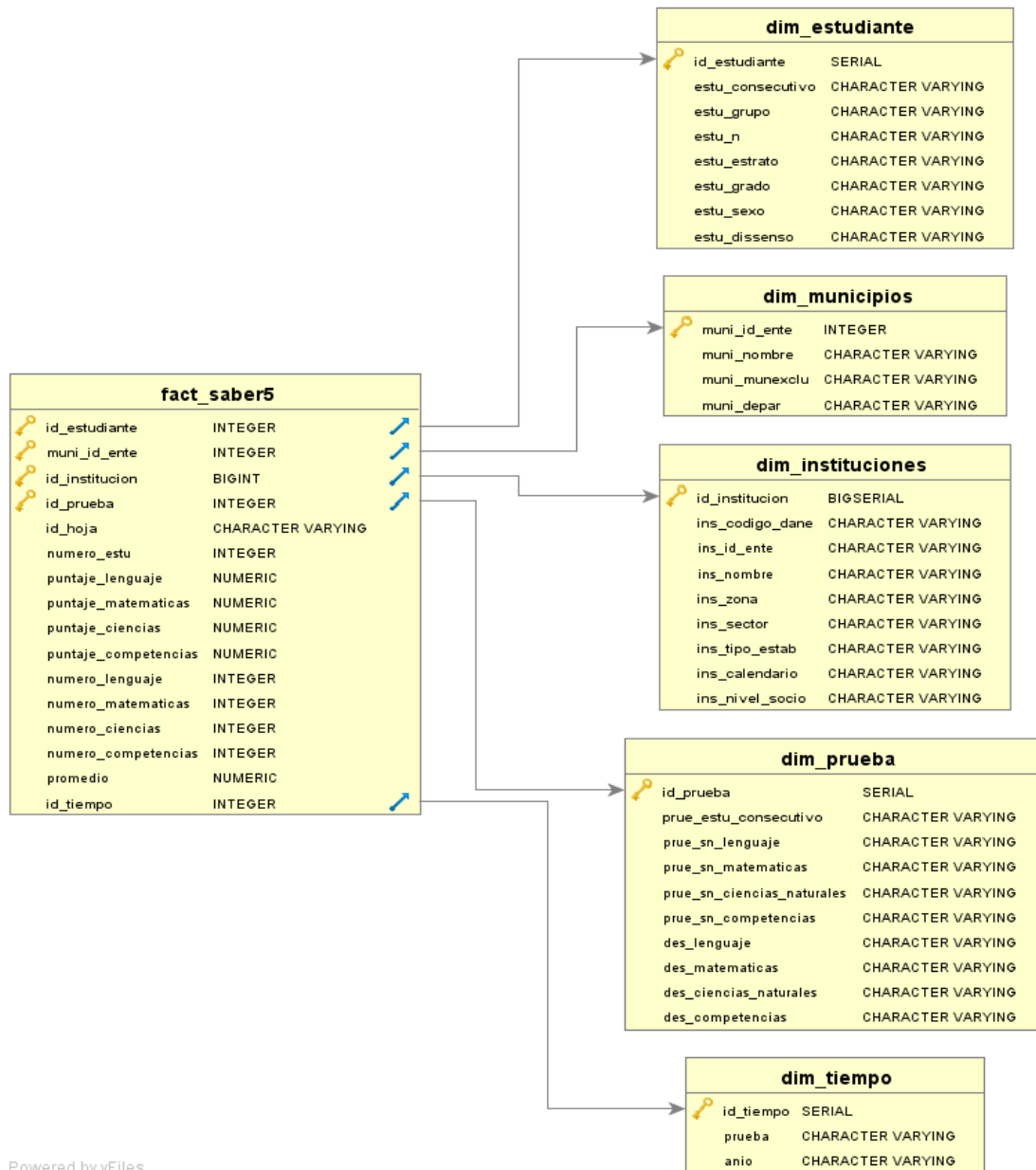
En la siguiente imagen podemos encontrar desplegado el directorio o carpeta template el cual contiene las diferentes plantillas bases que permite tener las diferentes rutas de librerías y JQuery que se utilizan para toda la aplicación con sus respectivas configuraciones, así también un pie de página y un encabezado como herencia para todas las plantillas que se visualicen para el usuario.



MODELADO DE DATOS

Para el diseño de la aplicación se decidió trabajar con diseño multidimensional con un esquema de estrella, este tipo de esquema de base de datos relacional consta de una sola tabla de hechos central rodeada de tablas de dimensiones.

En la siguiente Figura se muestra un esquema de estrella con una sola tabla de hechos y cuatro tablas de dimensiones. Un esquema de estrella puede tener cualquier número de tablas de dimensiones. Las ramas situadas al final de los enlaces que conectan las tablas indican una relación de muchos a uno entre la tabla de hechos y cada tabla de dimensiones con sus respectivas llaves primarias y llaves foráneas.



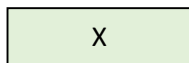
Powered by yFiles

MERCADO DE DATOS

En las siguientes tablas se presentan los atributos de los Valores plausibles, Entidades territoriales certificadas, Establecimientos y Municipios existentes en los repositorios que comprende los años de 2014 a 2017 que tienen datos indispensables para la agregación de la nueva bodega de datos.

En las siguientes tablas se analizan los atributos que están y no están presentes en cada uno de los repositorios.

Está presente



No está presente



Repositorio de datos (Valores plausibles).

VARIABLES	2014	2015	2016	2017	DESCRIPCIÓN DE VARIABLES
estu_consecutivo	X	X	X	X	ID público del estudiante
aplicación	X	X	X	X	No identificado
calendario	X	X	X	X	Calendario del establecimiento
departamento	X	X	X	X	Código DANE del departamento al que pertenece el establecimiento educativo
disenso	X	X	X	X	Marca de discapacidad cognitiva
enteterr	X	X	X	X	Código DANE de la entidad a la que pertenece el establecimiento educativo
establecimiento	X	X	X	X	Código DANE del establecimiento educativo al que pertenece el estudiante
estrato	X	X	X	X	Estrato al que pertenecen el establecimiento dentro del marco muestral
grado	X	X	X	X	Grado del estudiante

grupo	X	X	X	X	Grupo/salón al que pertenecen los estudiantes
id_hoja	X	X	X	X	Código de identificación de hoja de respuestas
jornada	X	X	X	X	Código de la sede jornada al que pertenece el estudiante
leng_copietas	X	X	X	X	Indicador de copia en lenguaje
leng_score1	X	X	X	X	Score1-valor plausible 1 de lenguaje
leng_score2	X	X	X	X	Score2-valor plausible 2 de lenguaje
leng_score3	X	X	X	X	Score3-valor plausible 3 de lenguaje
leng_score4	X	X	X	X	Score4-valor plausible 4 de lenguaje
leng_score5	X	X	X	X	Score5-valor plausible 5 de lenguaje
leng_weight	X	X	X	X	Peso muestral del lenguaje
mate_copietas	X	X	X	X	Indicador de copia en matemáticas
mate_score1	X	X	X	X	Score1-valor plausible 1 de matemáticas
mate_score2	X	X	X	X	Score2-valor plausible 2 de matemáticas
mate_score3	X	X	X	X	Score3-valor plausible 3 de matemáticas
mate_score4	X	X	X	X	Score4-valor plausible 4 de matemáticas
mate_score5	X	X	X	X	Score5-valor plausible 5 de matemáticas

mate_weight	X	X	X	X	Peso muestral de matemáticas
cien-copietas	X				Indicador de copia de ciencias
cien_score1	X				Score1-valor plausible 1 de ciencias
cien_score2	X				Score2-valor plausible 2 de ciencias
cien_score3	X				Score3-valor plausible 3 de ciencias
cien_score4	X				Score4-valor plausible 4 de ciencias
cien_score5	X				Score5-valor plausible 5 de ciencias
cien_weight	X				Peso muestral de ciencias
comp_copietas		X	X		Indicador de copia de competencias
comp_score1		X	X		Score1-valor plausible 1 de competencias
comp_score2		X	X		Score2-valor plausible 2 de competencias
comp_score3		X	X		Score3-valor plausible 3 de competencias
comp_score4		X	X		Score4-valor plausible 4 de competencias
comp_score5		X	X		Score5-valor plausible 5 de competencias
comp_weight		X	X		Peso muestral de competencias
modeloedu	X	X	X	X	No definido

municipio	X	X	X	X	Código DANE al que pertenece el establecimiento educativo
n	X	X	X	X	Matricula del grupo al que pertenecen
nivel	X	X	X	X	Nivel socioeconómico del establecimiento
sector	X	X	X	X	Sector del establecimiento
sexo	X	X	X	X	Sexo del estudiante
zona	X	X	X	X	Zona de la jornada
zonastab	X	X	X	X	Zona del establecimiento

Repositorio de datos (Entidades).

VARIABLES	2014	2015	2016	2017	DESCRIPCIÓN DE VARIABLES
id_ente	X	X	X	X	Identificador del ente territorial
nombre	X	X	X	X	Nombre del ente territorial
munexclu	X	X	X	X	Nombre de los municipios certificados excluidos dentro del ente territorial
tipo	X	X	X	X	Tipo de ente territorial 1=ETC, 2=DPTO,3= MPIO

Repositorio de datos (Establecimientos).

VARIABLES	2014	2015	2016	2017	DESCRIPCIÓN DE VARIABLES
cod_dane	X	X	X	X	Código DANE del establecimiento educativo
id_municipio	X	X	X	X	Código DANE del municipio al que pertenece el establecimiento educativo
id_ente	X	X	X	X	id de la entidad territorial
nombre	X	X	X	X	Nombre del establecimiento educativo reportado en el DUE
zona	X	X	X	X	Zona donde la mayoría de la población atendida por el establecimiento educativo se ubica 1=urbano 2=rural
sector	X	X	X	X	Naturaleza administrativa de establecimiento educativo según 1=oficial urbano 2=oficial rural 3=no oficial
tipo_estab	X	X	X	X	Tipo de establecimiento 1=Oficial urbano 2=Oficial rural 3=no oficial
calendario	X	X	X	X	Calendario del establecimiento
nivel_socio	X	X	X	X	NSE asignado de acuerdo a la clasificación realizada con puntajes por medios

Repositorio de datos (Municipios).

VARIABLES	2014	2015	2016	2017	DESCRIPCIÓN DE VARIABLES
id_ente	X	X	X	X	Identificador del ente territorial
nombre	X	X	X	X	Nombre del ente territorial
munexclu	X	X	X	X	Nombre de los municipios certificados excluidos dentro del ente territorial
tipo	X	X	X	X	Tipo del ente territorial posibles valores 3

ATRIBUTOS DE REPOSITORIO DE DATOS

Atributos (dimensión estudiante).

id_estudiante	integer
estu_consecutivo	character varying
estu_grupo	character varying
estu_n	character varying
estu_estrato	character varying
estu_grado	character varying
estu_sexo	character varying

estu_dissenso	character varying
---------------	-------------------

Atributos (dimensión instituciones).

id_institucion	integer
ins_codigo_dane	character varying
ins_id_ente	character varying
ins_nombre	character varying
ins_zona	character varying
ins_sector	character varying
ins_tipo_estab	character varying
ins_calendario	character varying
ins_nivel_socio	character varying

Atributos (dimensión municipios).

muni_id_ente	Integer
muni_nombre	character varying
muni_munexclu	character varying

muni_depar	character varying
------------	-------------------

Atributos (dimensión prueba).

id_prueba	Integer
prue_estu_consecutivo	character varying
prue_sn_lenguaje	character varying
prue_sn_matematicas	character varying
prue_sn_ciencias_naturales	character varying
prue_sn_competencias	character varying
des_lenguaje	character varying
des_matematicas	character varying
des_ciencias_naturales	character varying
des_competencias	character varying

Atributos (dimensión tiempo).

id_tiempo	Integer
prueba	character varying

anio	character varying
------	-------------------

Atributos (dimensión fact saber 5).

id_estudiante	Integer
muni_id_ente	Integer
id_institucion	Integer
id_prueba	Integer
id_hoja	character varying
numero_estu	Integer
puntaje_lenguaje	Numeric
puntaje_matematicas	Numeric
puntaje_ciencias	Numeric
puntaje_competencias	Numeric
numero_lenguaje	Integer
numero_matematicas	Integer
numero_ciencias	Integer
numero_competencias	Integer

promedio	Numeric
Id_tiempo	Integer

ANÁLISIS Y CALIDAD DE DATOS

Análisis calidad de datos de la bodega para identificar el número de registros nulos, no nulos y distintos cuenta con un total de **3’731.692** registros, se tuvo en cuenta las 5 dimensiones con el propósito de realizar un análisis de la calidad de los datos e identificar que los datos contenidos sean fiables

Análisis de calidad de datos (dimensión prueba).

No.	Atributos	No nulos	Nulos	Valores distintos	Valores distintos
1	id_prueba	13875	0	13875	"40"
					"41"
					"42"
					"43"
					"44"
2	prue_estu_consecutivo	13875	0	13875	"V201450046302"
					"V201450046303"
					"V201450046304"

					"V201450046306"
					"V201450046309"
3	prue_sn_lenguaje	9056	4819	2	"CON_COPIA"
					"NO_COPIA"
4	prue_sn_matematicas	9170	4705	2	"CON_COPIA"
					"NO_COPIA"
5	prue_sn_ciencias	6051	7824	2	"CON_COPIA"
					"NO_COPIA"
6	prue_sn_competencias	3011	10864	2	"CON_COPIA"
					"NO_COPIA"
7	des_lenguaje	9056	4819	0	'INSUFICIENTE'
					'MINIMO'
					'SATISFACTORIO'
					'AVANZADO'
8	des_matematicas	9170	4705	0	'INSUFICIENTE'
					'MINIMO'
					'SATISFACTORIO'
					'AVANZADO'
9	des_ciencias_naturales	6051	7824	0	'INSUFICIENTE'

10	des_competencias	3011	10864	0	'MINIMO'
					'SATISFACTORIO'
					'AVANZADO'
					'INSUFICIENTE'
					'MINIMO'
					'SATISFACTORIO'
					'AVANZADO'

Análisis de calidad de datos (dimensión municipio).

No.	Atributos	No nulos	Nulos	Valores distintos	Valores distintos
1	muni_id_ente	13	0	13	"52022"
					"52210"
					"522152"
					"52224"
					"52227"
2	muni_nombre	13	0	13	"ALDANA"
					"CONTADERO"
					"CORDOBA"
					"CUASPUD"

3	muni_munexclu	13	0	2	"CUMBAL"
					"ETC"
					"SIN_DATO"
4	muni_depar	13	0	1	"NARINO"

Análisis de calidad de datos (dimensión estudiante).

No.	Atributos	No nulos	Nulos	Valores distintos	Valores distintos
1	id_estudiante	13875	0	13875	"50"
					"51"
					"52"
					"53"
					"54"
2	estu_consecutivo	13875	0	13875	"V201450046302"
					"V201450046303"
					"V201450046304"
					"V201450046305"
					"V201450046306"
3	estu_grupo	13875	0	3	"01"
					"02"

					"99"
4	estu_n	13875	0	112	"1"
					"10"
					"100"
					"101"
					"103"
5	estu_estrato	13875	0	12	99E1209
					99E1219
					99E9999
					99O1109
					99O1209
6	estu_grado	13875	0	1	"5"
7	estu_sexo	13875	0	3	FEMENINO
					MASCULINO
					NO_ESPECIFICA
8	estu_disenso	13875	0	2	"N"
					"S"

Análisis de calidad de datos (dimensión institución).

No.	Atributos	No nulos	Nulos	Valores distintos	Valores distintos
1	id_institucion	358	0	358	"39"
					"40"
					"41"
					"42"
					"43"
2	ins_codigo_dane	358	0	358	"152022000084"
					"152210000261"
					"152215000138"
					"152224000019"
					"152227000028"
3	ins_id_ente	358	0	2	48
					89
4	ins_nombre	358	0	336	-CENTRO
					EDUCATIVO SAN MARTIN
					CENTRO EDUCATIVO

					CASAFRIA
					CENTRO EDUCATIVO ARVELA
					CENTRO EDUCATIVO CUAS
					CENTRO EDUCATIVO GUAN
5	ins_zona	358	0	2	RURAL URBANO
6	ins_sector	358	0	2	NO_OFICIAL OFICIAL
7	ins_tipo_estab	358	0	3	NO_OFICIAL OFICIAL_RURAL OFICIAL_URBANO
8	ins_calendario	358	0	1	A
9	ins-nivel_socio	358	0	5	1 2

	3
	4
	5

Análisis de calidad de datos (dimensión tiempo).

No.	Atributos	No nulos	Nulos	Valores distintos	Valores distintos
1	Id_tiempo	13875	0	0	"52022"
					"52210"
					"522152"
					"52224"
					"52227"
2	prueba	13875	0	0	"V201450046302"
					"V201450046303"
					"V201450046304"
					"V201450046305"
					"V201450046306"
3	anio	13875	0	0	"2014"
					"2015"
					"2016"

	"2017"
--	--------

Análisis de calidad de datos (dimensión fact_saber5).

No.	Atributos	No nulos	Nulos	Valores distintos	Valores distintos
1	id_estudiante	13875	0	13875	"35"
					"36"
					"37"
					"38"
					"39"
2	muni_id_ente	13875	0	13	"52022"
					"52210"
					"52215"
					"52224"
					"52227"
3	id_institucion	13875	0	324	"150"
					"151"
					"163"
					"167"
					"168"

4	id_prueba	13875	0	13875	"1"
					"55"
					"60"
					"70"
					"100"
5	id_hoja	13875	0	13875	"V201450046302"
6	numero_estu	13875	0	1	1
7	puntaje_lenguaje	9056	4819	151	"1"
					"1,2"
					"1,25"
					"1,33"
					"1,38"
8	puntaje_matematicas	9170	4705	154	"146"
					"152"
					"152,67"
					"153"
					"154"
9	puntaje_ciencias	6051	7824	102	"1,45"
					"146"
					"147"

10	puntaje_competencias	3011	10864	72	"148"
					"148,78"
					"150,104"
					"150,943"
					"152"
11	numero_lenguaje	13875	0	2	"163"
					"0"
12	numero_matematicas	13875	0	2	"1"
					"0"
13	numero_ciencias	13875	0	2	"1"
					"0"
14	numero_competencias	13875	0	2	"1"
					"0"
15	promedio	13875	0	0	"1"
					"146"
					"147"
					"148"

16	Id_tiempo	13875	0	0	"1"
					"55"
					"60"
					"70"
					"100"