

Data Visualization and Data Analysis

' Palíndromo

25

'Palíndromo

nombre masculino

Significado. Palabra o expresión que se lee de igual manera de izquierda a derecha y viceversa.

Ej. Anita lava la tina.

Ej. Amo la pacífica paloma.

Ej. Oír a diario.

Ej. Oirás orar a Rosario.

Ej. Yo hago yoga hoy.

Ej. Sé verla al revés.

Ej. Ojo rojo.

Ej. Dabale arroz a la zorra el abad



25 años de edad

Datos: ¿Qué son?

¿Qué son los datos?

- La importancia de la **semántica** de los datos.
- Una lista de números: 14, 2.6, 30, 30, 15, 100001, sin un significado o contexto (su semántica), son sólo números.
- Necesitamos información del significado de esos números, en la vida real.

Tablas de datos

M	2	C	C	37	59	161	160
V	1	O	O	41	62	178	181
M	2	O	O	36	50	159	153
V	2	O	O	42	69	176	179
V	2	O	O	43	74	175	179
M	3	C	C	37	62	169	165
M	2	O	O	36	56	162	158
M	2	O	O	37	58	162	163

Datos: ¿Qué son?

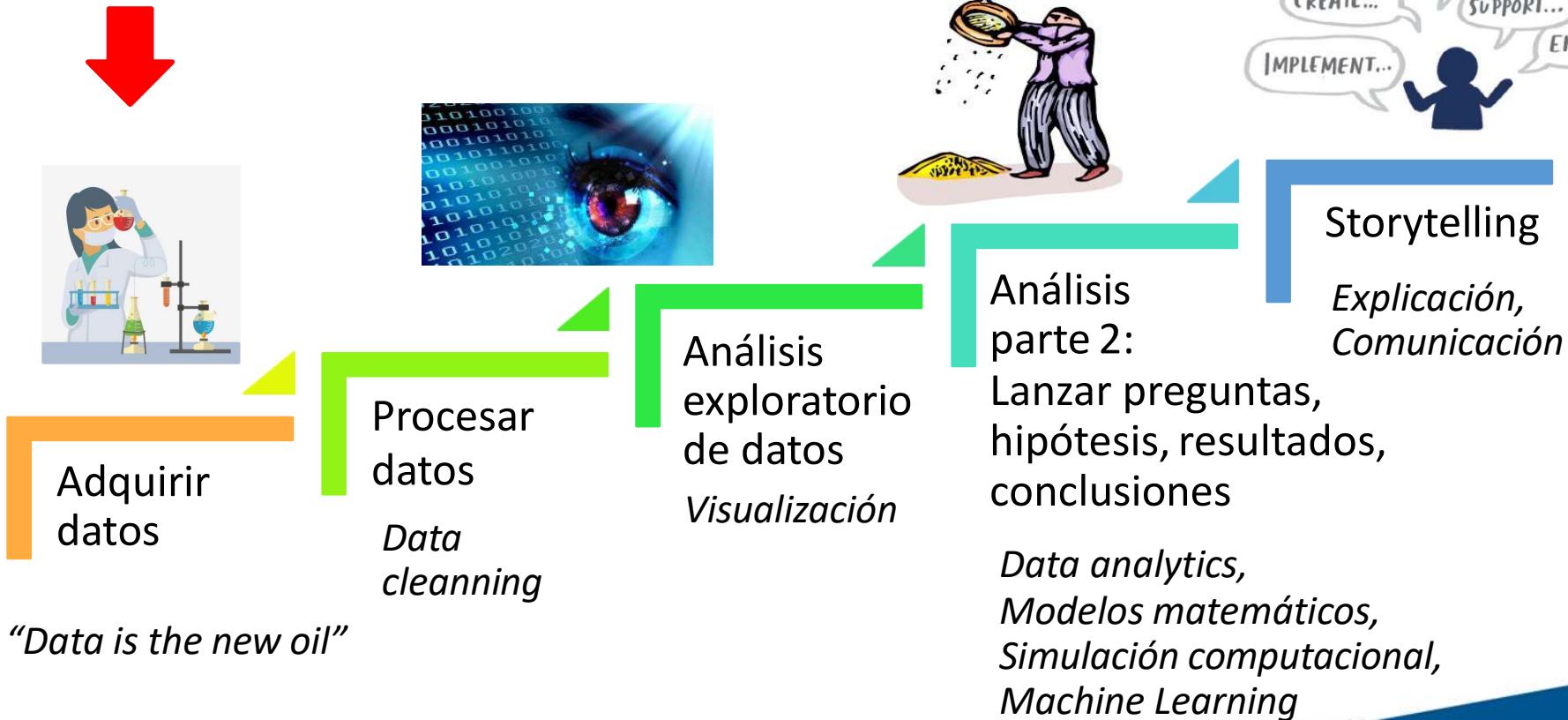
Tablas de datos

En ocasiones las tablas de datos proveen la semántica necesaria de los datos. Por ejemplo, a través de los nombres de las columnas, que dan el significado de los atributos.

Nombres de las columnas

Sexo	Deporte	Ojos	Pelo	N. calzado	Peso (Kg.)	Talla (cm.)	L. brazos (cm.)
M	2	C	C	37	59	161	160
V	1	O	O	41	62	178	181
M	2	O	O	36	50	159	153
V	2	O	O	42	69	176	179
V	2	O	O	43	74	175	179
M	3	C	C	37	62	169	165
M	2	O	O	36	56	162	158
M	2	O	O	37	58	162	163

El proceso de análisis y visualización de datos



Datos: Adquiriendo, procesando y analizando

Adquiriendo datos

A partir de:

- Experimentos, sensores, dispositivos, encuestas, simulaciones computacionales.
- Aplicaciones y repositorios web.
- Expertos en el área.



Datos: Adquiriendo, procesando y analizando

Adquiriendo datos

Formatos del archivo de datos:

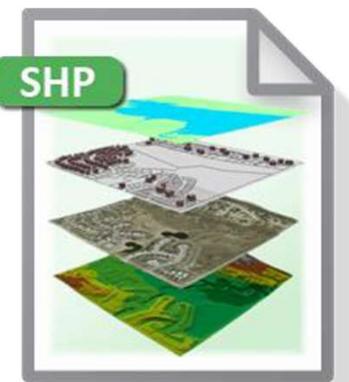
- Archivo de valores separados por comas:
comma separated values (csv).
(Representar datos en forma de tabla, en las que las columnas se separan por comas (o punto y coma en donde la coma es el separador decimal)
- Excel (xls, xlsx)
- JavaScript Object Notation (JSON)
- Extensible Markup Language (XML)
- Geographic Information Systems (GIS)

Sexo	Deporte	Ojos	Pelo	N. calzado	Peso (Kg.)	Talla (cm.)	L. brazos (cm.)
M	2	C	C	37	59	161	160
V	1	O	O	41	62	178	181
M	2	O	O	36	50	159	153
V	2	O	O	42	69	176	179
V	2	O	O	43	74	175	179
M	3	C	C	37	62	169	165
M	2	O	O	36	56	162	158
M	2	O	O	37	58	162	163

```

1  {
2   "users": [
3     {
4       "name": "John",
5       "age": 25
6     },
7     {
8       "name": "Mark",
9       "age": 29
10    },
11    {
12      "name": "Sarah",
13      "age": 22
14    }
15  ],
16  "dataTitle": "JSON Tutorial!",
17  "swiftVersion": 2.1
18 ]

```



Datos: Adquiriendo, procesando y analizando

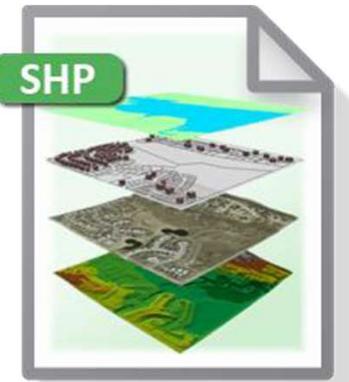
Adquiriendo datos

Formatos del archivo de datos:

- Archivo de valores separados por comas:
comma separated values (csv)
- Excel (*xls*, *xlsx*)
- JavaScript Object Notation (*JSON*)
- Extensible Markup Language (*XML*)
- Geographic Information Systems (*G/S*)

Sexo	Deporte	Ojos	Pelo	N. calzado	Peso (Kg.)	Talla (cm.)	L. brazos (cm.)
M	2	C	C	37	59	161	160
V	1	O	O	41	62	178	181
M	2	O	O	36	50	159	153
V	2	O	O	42	69	176	179
V	2	O	O	43	74	175	179
M	3	C	C	37	62	169	165
M	2	O	O	36	56	162	158
M	2	O	O	37	58	162	163

```
1 {  
2   "users": [  
3     {  
4       "name": "John",  
5       "age": 25  
6     },  
7     {  
8       "name": "Mark",  
9       "age": 29  
10    },  
11    {  
12      "name": "Sarah",  
13      "age": 22  
14    }  
15  ],  
16  "dataTitle": "JSON Tutorial!",  
17  "swiftVersion": 2.1  
18 }
```



**** No olvidar ajustar la configuración para que los decimales sean registrados por un “punto” y los miles por “coma”.**

Autoguardado

Libro1 - Excel

Buscar

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Automatizar Ayuda

Comentarios Compartir

Pegar Portapapeles Fuente Alineación Número Estilos Celdas Edición Confidencialidad Complementos Analizar Copilot datos

A1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1															
2															
3															
4															
5															
6															
7															
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															

Hoja1 +

Lista Accesibilidad: todo correcto

Buscar

3

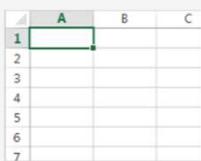
ESPLAA 11:52 a.m. 22/07/2025



Inicio

Nueva

Más plantillas →



Libro en blanco



Programación semanal de t...



Programación para los alu...



Diagrama de Gantt ágil



Le damos la bienvenida a E...



Buscar

Recientes

Favoritos

Compartidos conmigo

Nombre

Fecha de modificación



estado-actual-estudiantes-26-06-2025

OneDrive de Andrés Franco Herrera » Documentos » doctorado-CAyS-2025 » estudiantes-status-2025

Hace 1 h



2-flujo-de-caja-mensual-2025

OneDrive de Andrés Franco Herrera » Documentos » finanzas-personales-2025 » 2-flujo-de-caja-2025

Hace 1 h



plan-de-pagos-inversionistas-opcion-1

OneDrive de Andrés Franco Herrera » Documentos » finanzas-personales-2025 » 14b-proyecto-monterra-2025

Ayer a las 4:29 p. m.



plan-de-pagos-inversionistas-opcion-2

OneDrive de Andrés Franco Herrera » Documentos » finanzas-personales-2025 » 14b-proyecto-monterra-2025

Ayer a las 4:29 p. m.



Cuenta

Opciones



Buscar



ESP

LAA

11:54 a. m.
22/07/2025

Autoguardado

Archivo Inicio Insertar Disposito

Pegar Portapapeles A1

A B C

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Hoja1 +

Lista Accesibilidad: todo correcto

Buscar Buscar

3

ESP LAA 11:54 a.m. 22/07/2025

Opciones de Excel

General

Fórmulas

Datos

Revisión

Copilot

Guardar

Idioma

Accesibilidad

Avanzadas

Personalizar cinta de opciones

Barra de herramientas de acceso rápido

Complementos

Centro de confianza

Opciones generales para trabajar con Excel.

Opciones de interfaz de usuario

Cuando se utilizan varias pantallas:

- Optimizar la apariencia
- Optimizar para compatibilidad (requiere reiniciar la aplicación)

Mostrar minibarra de herramientas al seleccionar

Mostrar opciones de análisis rápido durante la selección

Mostrar Convertir a Tipos de datos al escribir

Habilitar vista previa activa

Contraer la cinta de opciones automáticamente

Contraer el cuadro de Búsqueda de Microsoft de manera predeterminada

Estilo de información en pantalla: Mostrar descripciones de características en información en pantalla

Al crear nuevos libros

Usar esta fuente como fuente predeterminada: Fuente de cuerpo

Tamaño de fuente: 11

Vista predeterminada para hojas nuevas: Vista normal

Incluir este número de hojas: 1

Personalizar la copia de Microsoft Office

Nombre de usuario: evaluador externo

Usar siempre estos valores sin tener en cuenta el inicio de sesión en Office.

Aceptar Cancelar

Comentarios Compartir

Analizar Copilot datos

N O

Archivo Inicio Insertar Dispositos

Pegar Portapapeles A1 fx

A B C

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Hoja1 +

Lista Accesibilidad: todo correcto

Buscar Buscar

Autoguardado

Opciones de Excel

General Fórmulas Datos Revisión Copilot Guardar Idioma Accesibilidad Avanzadas

Personalizar cinta de opciones Barra de herramientas de acceso rápido Complementos Centro de confianza

Opciones avanzadas para trabajar con Excel.

Opciones de edición

Despues de presionar Entrar, mover selección Dirección: Hacia abajo

Insertar automáticamente un punto decimal Posiciones: 2

Permitir arrastrar y colocar el controlador de relleno y las celdas Mensaje de alerta antes de reemplazar celdas

Permitir editar directamente en las celdas Extender formatos de rangos de datos y fórmulas

Habilitar la inserción automática de porcentajes

Habilitar Autocompletar para valores de celda Relleno rápido automático

Hacer zoom al usar la rueda de IntelliMouse

Avisar al usuario cuando haya una operación que requiera mucho tiempo Cuando este número de celdas (en miles) se vean afectadas: 33.554

Usar separadores del sistema Separador decimal: . Separador de miles: ,

Movimiento del cursor:

Lógico Visual

Aceptar Cancelar

Comentarios Compartir

Analizar Copilot datos

N O

ESP LAA 11:55 a.m. 22/07/2025

Datos: Adquiriendo, procesando y analizando

Adquiriendo datos

Ejemplo:

Datos del IDEAM.

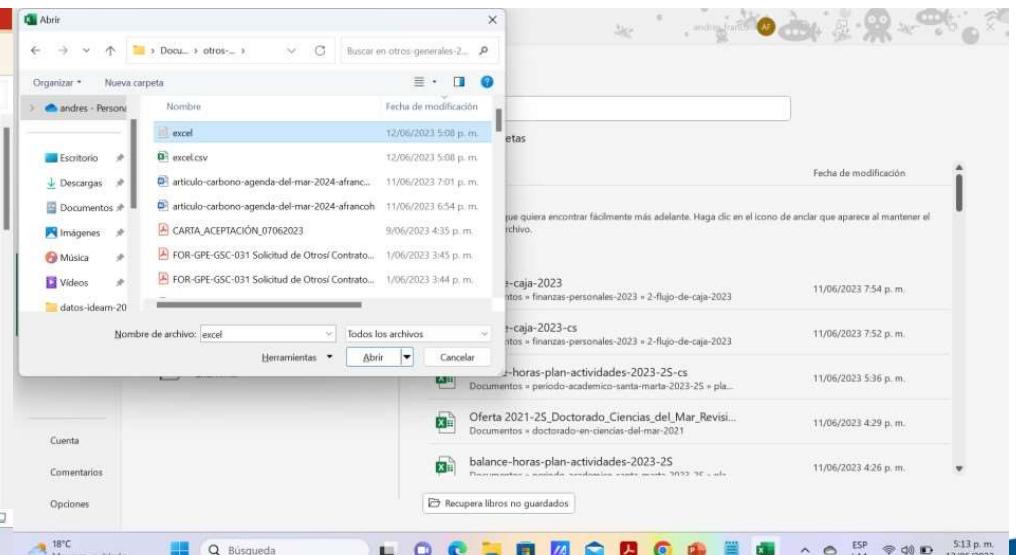
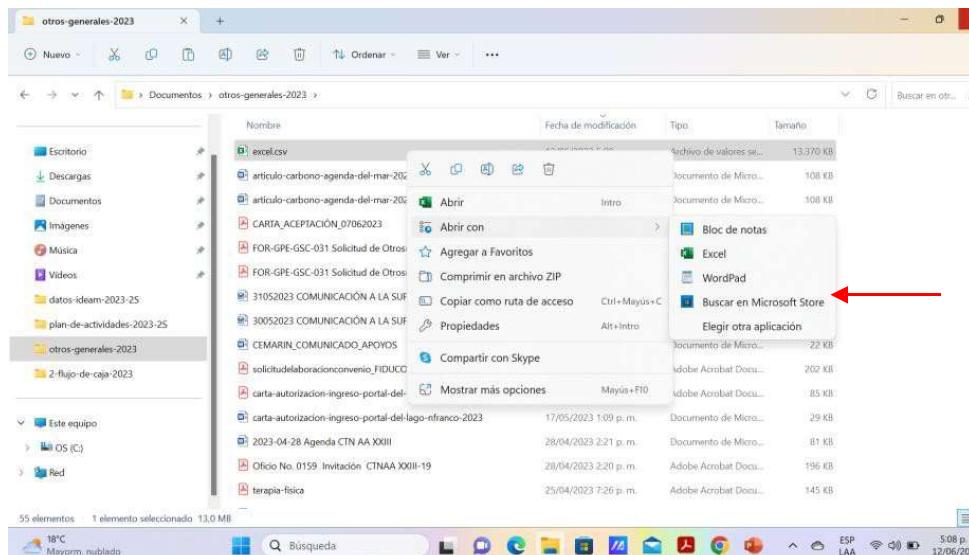


Datos: Adquiriendo, procesando y analizando



Archivo .csv

Transformar a .txt



**** No olvidar ajustar la configuración para que los decimales sean registrados por un “punto” y los miles por “coma”.**

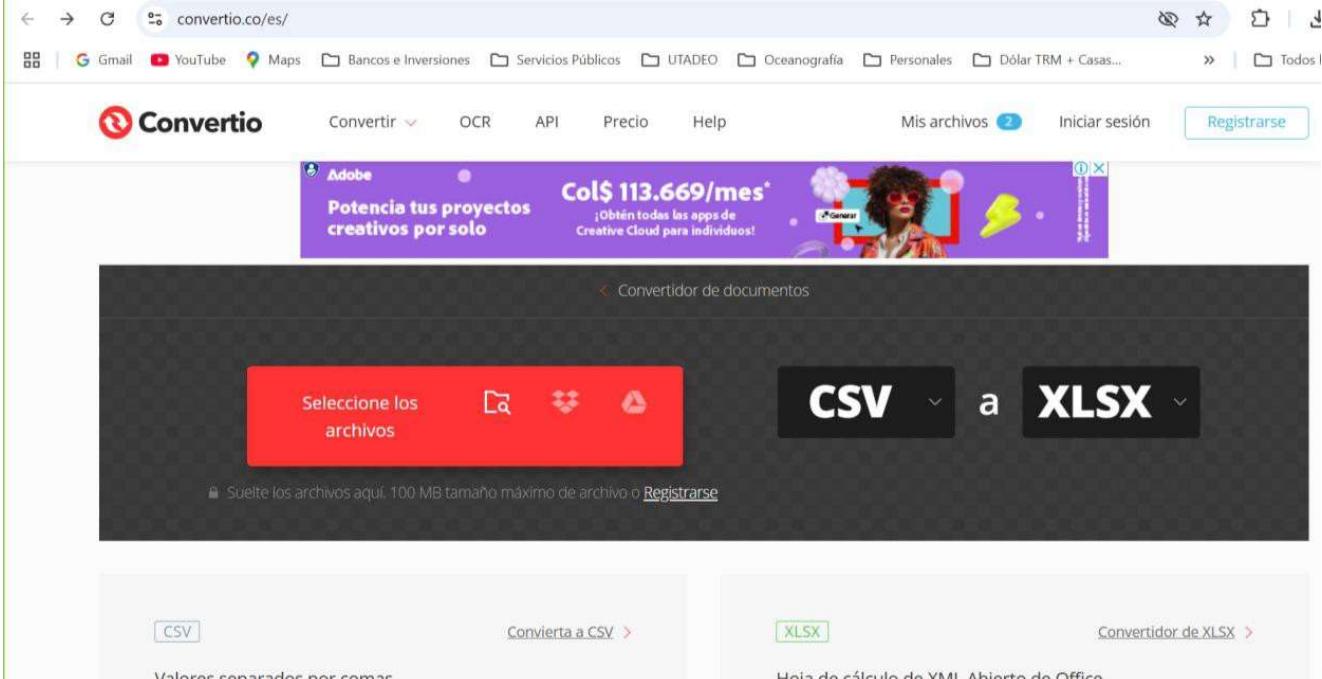
Datos: Adquiriendo, procesando y analizando



Archivo .csv

https://www.youtube.com/shorts/vbo_9WJML1Q

<https://convertio.co/es/>



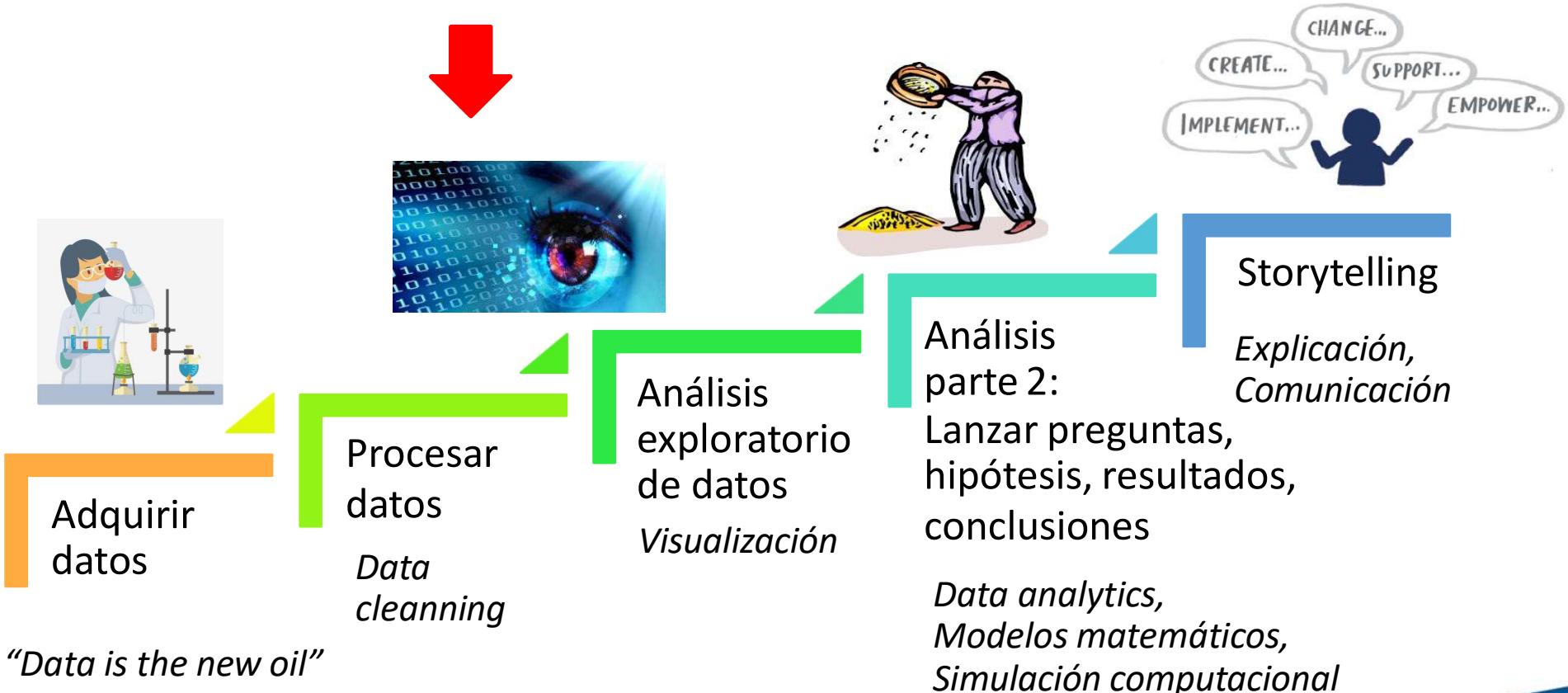
The screenshot shows the homepage of convertio.co/es/. At the top, there's a navigation bar with links like 'Convertir', 'OCR', 'API', 'Precio', 'Help', 'Mis archivos', 'Iniciar sesión', and 'Registrarse'. Below the navigation is a banner for Adobe Creative Cloud with the text 'Potencia tus proyectos creativos por solo Col\$ 113.669/mes*'. The main area features a large red button labeled 'Seleccione los archivos' with three file icon options (CSV, XLS, XLSX). To the right, there are dropdown menus for 'CSV' and 'a' followed by 'XLSX'. Below these are two smaller sections: 'Valores separados por comas' under 'CSV' and 'Hoja de cálculo de XML Abierto de Office' under 'XLSX'. A dashed arrow points from the 'Archivo .csv' icon on the left to the 'CSV' section of the website.



Archivo .xlsx

**** No olvidar ajustar la configuración para que los decimales sean registrados por un “punto” y los miles por “coma”.**

El proceso de análisis y visualización de datos



Datos: Adquiriendo, procesando y analizando

Procesando los datos: Tablas de datos

Registro o ítem

Es una entidad individual, discreta.

Ejemplo: Personas, orden de compra, tiendas, países.

Variable o atributo (feature)

Alguna propiedad específica que puede ser medida, observada o registrada.

Ejemplo: Salario, precio, ventas mensuales, pobreza.

Tabla. Ejemplo de datos obtenidos de la encuesta aplicada a una clase de 60 estudiantes en una universidad en España. Tomado de: [*Estadística con proyectos*](#), Carmen Batanero y Carmen Díaz (editoras)

		Atributo					
Sexo	Deporte	Ojos	Pelo	N. calzado	Peso (Kg.)	Talla (cm.)	L. brazos (cm.)
M	2	C	C	37	59	161	160
V	1	O	O	41	62	178	181
M	2	O	O	36	50	159	153
V	2	O	O	42	69	176	179
V	2	O	O	43	74	175	179
M	3	C	C	37	62	169	165
M	2	O	O	36	56	162	158
M	2	O	O	37	58	162	163

Item

Celda

Datos: Adquiriendo, procesando y analizando

Procesando los datos: Tipos de variables



Tabla 1.2. Tipos de variables

		Los valores que toma la variable:
Cualitativa	Ordinal	Se pueden clasificar en grupos ordenados
	Nominal	Se pueden clasificar en grupos pero no hay un orden entre ellos
Cuantitativa	Discreta	Son finitos o contables
	Continua	Son infinitos en cualquier intervalo razonable

Fuente: elaboración propia.

Datos: Adquiriendo, procesando y analizando

Procesando los datos: Tipos de variables

Variable cualitativa ordinal: identifica características que se pueden clasificar en grupos que admiten orden. Ejemplos de esto son el resultado de una entrevista (rechazado, normal y sobresaliente) o el premio en una prueba olímpica (oro, plata, bronce).

Variable cualitativa nominal: identifica características que se pueden clasificar en grupos que difieren o son iguales en una propiedad, por ejemplo, el municipio de nacimiento o el color de los ojos. Observe que en las variables cualitativas nominales las características no admiten un orden natural.

Datos: Adquiriendo, procesando y analizando

Procesando los datos: Tipos de variables

Variable cuantitativa discreta: es aquella que toma un número finito o contable^{1*} de valores. El número de respuestas correctas en un examen de cinco preguntas {0,1,2,3,4,5} o el número de hijos de una persona {0,1,2,...} son ejemplos de estas variables.

Variable cuantitativa continua: es aquella que puede tomar infinitos valores en cualquier intervalo razonable de la variable. La estatura de un hombre adulto es una variable aleatoria continua. Note que existe una infinidad de posibilidades para la estatura en un intervalo arbitrario, por ejemplo, entre los 160 y 161 centímetros. Aunque las personas suelen responder con un número entero cuando se les pregunta por su estatura (en centímetros), esta variable se puede medir con la precisión deseada utilizando el instrumento adecuado.

Datos: Adquiriendo, procesando y analizando



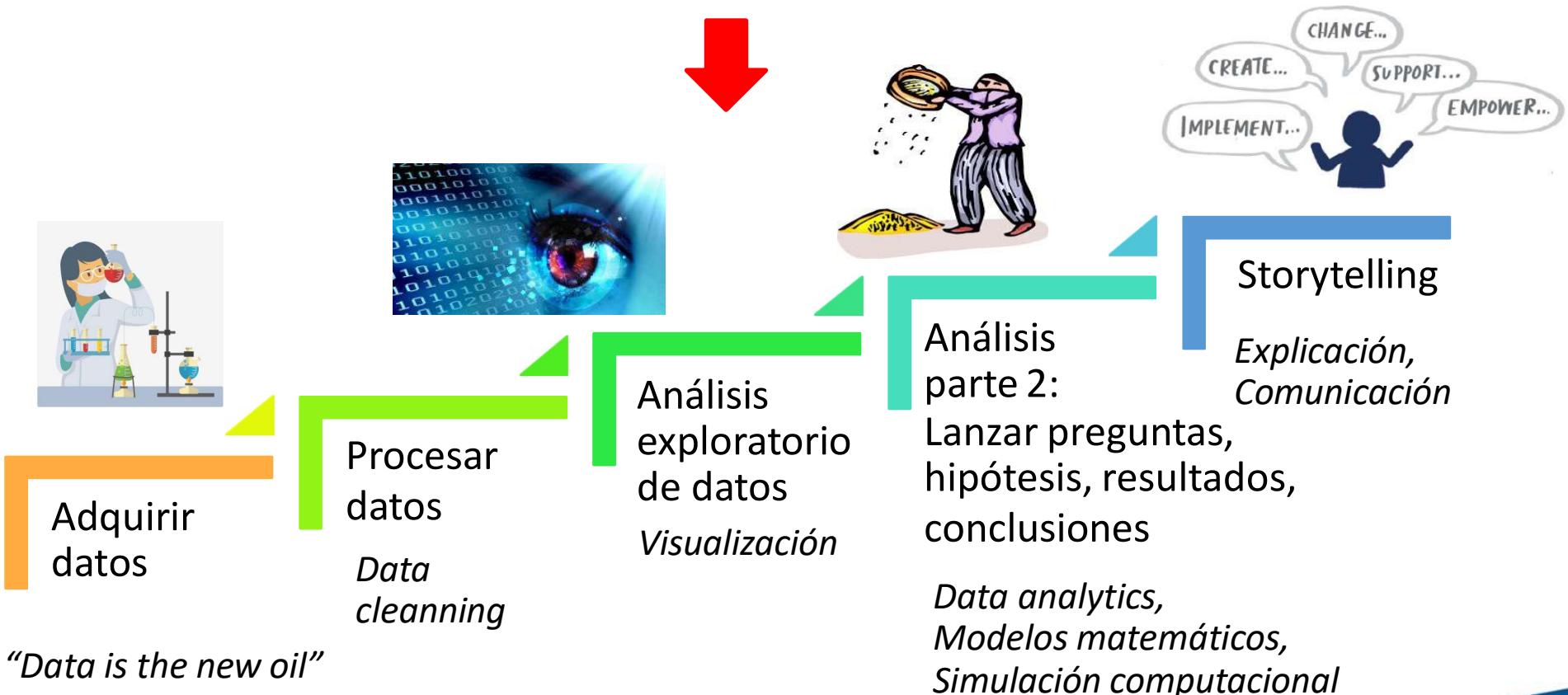
Limpiar y Ajustar
[Eliminar Información no relevante]

Abrir los siguientes archivos

**2b. Vientos-aeropuerto-manizales-2020-
2021.csv**

**2c. Matriz de dirección de viento febrero
2020**

El proceso de análisis y visualización de datos



Datos: Adquiriendo, procesando y analizando

Analizando datos cualitativos: tablas de frecuencia y gráficos barras

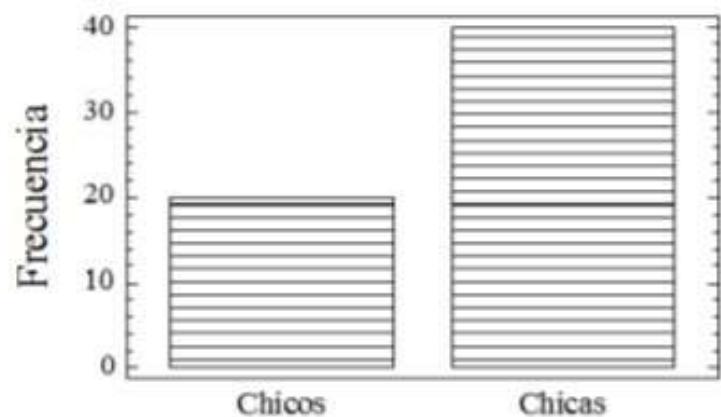
Tabla 3.1. Distribución de frecuencia del sexo de los alumnos

Sexo	Frecuencia absoluta	Frecuencia relativa	Porcentaje
Chicos	23	0,3833	38.3
Chicas	37	0,6167	61.7
Total	60	1	100

$$\text{Frecuencia relativa}(i) = \frac{\text{Frecuencia absoluta } (i)}{\text{Total}}$$

$$\text{Porcentaje}(i) = \frac{\text{Frecuencia absoluta } (i)}{\text{Total}} \times 100$$

Figura 3.1. Distribución de alumnos por sexo



[Estadística con proyectos](#), Carmen Batanero y Carmen Díaz (editoras)

Datos: Adquiriendo, procesando y analizando

Analizando datos cualitativos: tablas de frecuencia y gráficos barras

Figura 3.2. Distribución de alumnos, según práctica de deporte

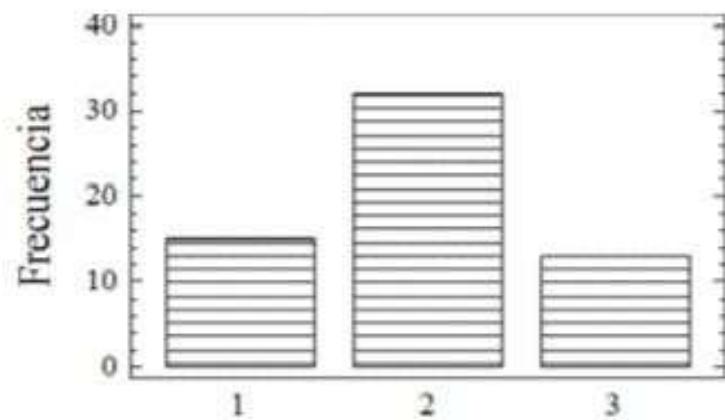
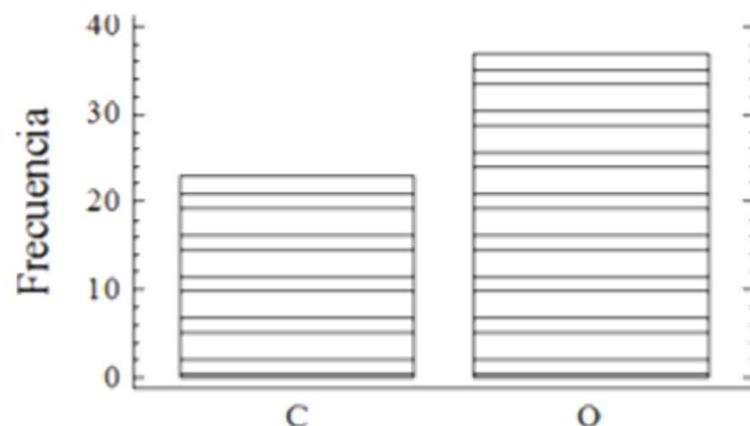


Figura 3.3. Distribución de alumnos según el color de los ojos



[Estadística con proyectos](#), Carmen Batanero y Carmen Díaz (editoras)

Datos: Adquiriendo, procesando y analizando



Crear Filtros
+
Figura Básica

Abrir Archivo
**3-Matriz Ejemplo Tabla Exploratoria
y Filtros.xlsx**

El proceso de análisis y visualización de datos

Taller: Características de un estudiante de Data Visualization

Objetivo: Elaborar un perfil de los estudiantes de la clase, identificando el estudiante típico y analizando si hay diferencias entre el estudiante y la estudiante típicos, respecto a sus características físicas, edad, semestre y lugar de residencia.

Pregunta: ¿Cómo son los estudiantes de Data Viz, en cuanto a sus características físicas, edad, semestre y lugar de residencia?

Herramientas: Se realizarán tablas de frecuencia, de contingencia y visualizaciones básicas (barras) utilizando software (Excel).

Datos: Adquiriendo, procesando y analizando

Adquiriendo datos

Taller: Características de un estudiante de Data Visualization

Paso 1: Cada estudiante del semestre 2026-1S diligenciará una encuesta sobre *¿Cómo son los estudiantes de Data Viz?*

https://forms.office.com/Pages/DesignPageV2.aspx?prevorigin=shell&origin=NeoPortaIPage&subpage=design&id=-4HJLoOSoEOaOTMy_FTxDLEeD2Xq4sIPtAOiqb9D1cpUQ0hLMzI5QzhPM0FIOFhFUEhDVEJRV1VUNS4u&analysis=false

Datos: Adquiriendo, procesando y analizando

Adquiriendo datos

Taller: Características de un estudiante de Data Visualization

Paso 1: Cada estudiante del semestre 2026-1S diligenciará una encuesta sobre *¿Cómo son los estudiantes de Data Viz?*

Encuesta Data Viz



Datos: Adquiriendo, procesando y analizando

Procesando los datos: Tipos de variables

Taller: Características de un estudiante de Data Visualization

Paso 2: Identifique el número de registros y el nombre y tipo de cada variable.

Datos: Adquiriendo, procesando y analizando

Procesando los datos: operaciones

En Excel o Google Sheets

- Filtros para **ordenar** (Excel, Google Sheets).

Taller: Características de un estudiante de Data Visualization

Paso 3: Cree *filtros* en la tabla de datos y ordene de acuerdo a alguna variable.

Datos: Adquiriendo, procesando y analizando

Analizando datos cualitativos: tablas de contingencia

Tabla 3.4. Clasificación cruzada de alumnos según color de pelo y ojos

	Ojos claros	Ojos oscuros	Total
Pelo claro	16	11	27
Pelo oscuro	5	28	33
Total	21	39	60

[Estadística con proyectos](#), Carmen Batanero y Carmen Díaz (editoras)

Tabla 3.5. Distribución condicional de color de ojos según color de pelo

	Ojos claros	Ojos oscuros	Total
Pelo claro	16	11	27
	59 %	41%	
Pelo oscuro	5	28	33
	15%	85%	
Total	21	39	60
	35%	65%	

Taller: Características de un estudiante de Data Visualization

Paso 4: Visualice mediante gráficos de barras el conteo (frecuencias absolutas) y porcentajes (frecuencias relativas) de variables cualitativas.

1. Entregar en la plantilla disponible en AVATA (semana 2; plantilla-taller-estudiantes-viz-2026-1S.xlsx) el taller que contenga:

- i. Matriz de base de datos de la encuesta en .xlsx limpia y organizada.
- ii. Matriz de base de datos de la encuesta en .xlsx limpia y organizada, con filtros y dos variables filtradas.
- iii. Generar una tabla de frecuencias absolutas y frecuencias relativas con la variable que a bien considere de las que están disponibles en la tabla.
- iv. Enviar al correo electrónico hugo.escobar@utadeo.edu.co en las siguientes fechas:

Grupo de los lunes (sábado 31 de enero). Máximo hasta las 6:00 p.m.

Grupo de los martes (Domingo 1 de febrero) Máximo hasta las 6:00 p.m.

Grupo de los miércoles (Lunes 2 de febrero) Máximo hasta las 6:00 p.m.

2. Libro Análisis exploratorio de Datos de Leandro González: leer las siguientes secciones (páginas 15 a 41):

- ¿Qué es la estadística?
- Población, muestra y variables.
- Gráficos para resumir datos cualitativos.

3. Ver los Videos

- [La belleza de la visualización de datos \(David McCandless\)](#)
- [Identificar una muestra y una población \(2:18 min\)](#)

4. Alistarse para el Quiz No. 2. Sobre los temas visto en esta clase.

Datos: Adquiriendo, procesando y analizando

Adquiriendo datos: repositorios

Buscadores generales: Google, [WolframAlpha](#), Our World In Data.

Fuentes de datos generales: [Kaggle](#), [Wikidata](#), [Amazon Public Data Sets](#).

Por temas:

Académicos: [Academic Torrents](#)

Deportes: [Basketball Reference](#), [Soccer Kaggle](#)

El mundo: [Our World in Data](#), [UNdata](#), [Organización Mundial de la Salud](#), [OECD](#),
[Banco Mundial](#)

Gobierno y política: [Data.gov](#), [Follow the money](#)

Colombia: [Terridata](#), [Datos Abiertos](#), [Instituto Nacional de Salud](#)

Ciencia: [USGS](#), [Environmental Data Initiative](#), [CODATA](#)

Datos: Adquiriendo, procesando y analizando

Adquiriendo datos: repositorios

Recurso	Página Web
Gapminder	https://www.gapminder.org/
Kaggle	https://www.kaggle.com/
Our World In Data	https://ourworldindata.org/
Víctimas de Conflicto armado en Colombia de la Unidad de Víctimas	https://cifras.unidadvictimas.gov.co/
Encuesta de Consumo Cultural del DANE	https://www.dane.gov.co/index.php/estadisticas-por-tema/cultura/consumo-cultural
Departamento Nacional de Planeación - Terridata	https://territoria.dnp.gov.co/
Datos Abiertos Gobierno de Colombia	https://www.datos.gov.co/
Instituto Nacional de Salud	https://www.ins.gov.co/Paginas/Inicio.aspx
United Nations Data	http://data.un.org/
Organización Mundial de la Salud	https://www.who.int/data/
OECD	https://stats.oecd.org/
Banco Mundial	https://data.worldbank.org/
Wikidata	https://www.wikidata.org/wiki/Wikidata:Main_Page
Amazon Public Data Set	https://aws.amazon.com/opendata/open-data-sponsorship-program/
U.S. Government's open data	https://www.data.gov/
U.S. Geological Survey	https://www.usgs.gov/products/data-and-tools/overview
Environmental Data Initiative	https://portal.edirepository.org/nis/home.jsp