

PRÁCTICA 09

ALUMNOS:

->CALDERÓN FERNÁNDEZ GABRIEL

->FLORES GONZÁLEZ LUIS BRANDON

->SANTAELLA MARÍN HÉCTOR

CONJUNTO DE DATOS

El dataset que tiene por nombre Adult.cvs tiene como objetivo la predicción para determinar si una persona gana más de 50 mil al año. Este origen de datos fue encontrado en

<https://archive.ics.uci.edu/ml/datasets/Adult>.

Para ser más exactos, las características de este conjunto son:

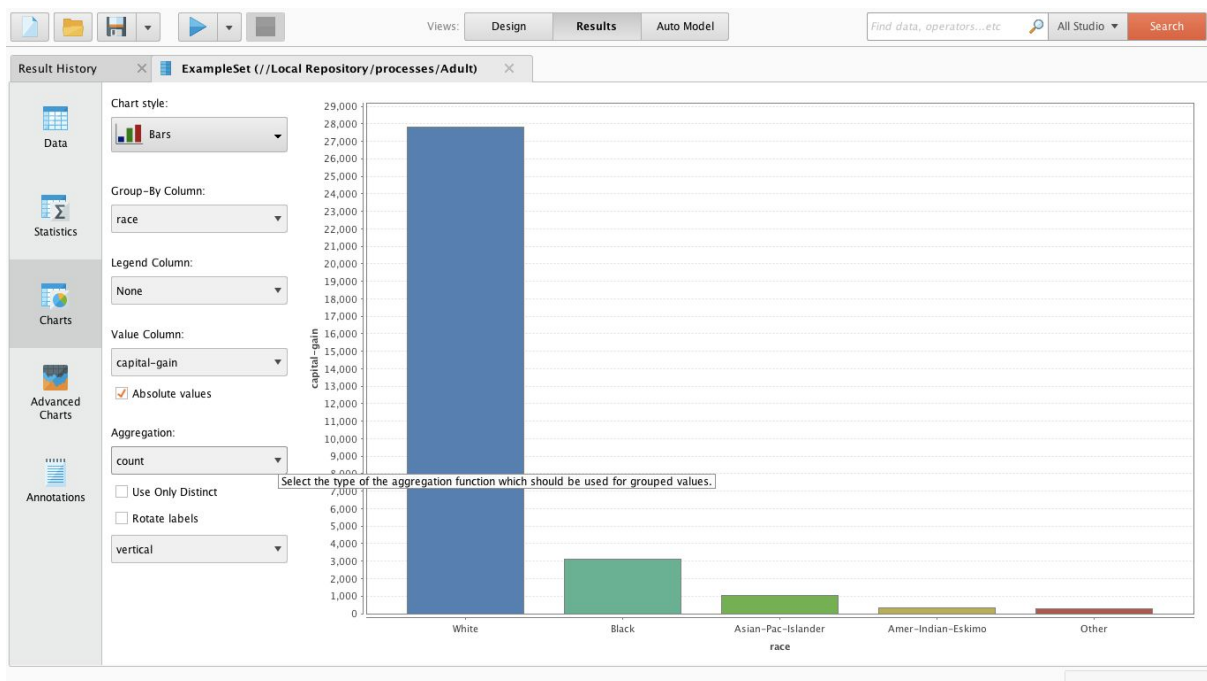
Data Set Characteristics:	Multivariate	Number of Instances:	48842	Area:	Social
Attribute Characteristics:	Categorical, Integer	Number of Attributes:	14	Date Donated	1996-05-01
Associated Tasks:	Classification	Missing Values?	Yes	Number of Web Hits:	1141948

TABLA DE DATOS

Views: Design Results Auto Model													
Find data, operators...etc													
All Studio Search													
Result History ExampleSet (/Local Repository/processes/Adult)													
ExampleSet (32560 examples, 0 special attributes, 15 regular attributes)													
Filter (32,560 / 32,560 examples): all													
Row No.	age	Work class	fnlwtg	Education	education-...	marital-sta...	occupation	relationship	race	sex	capital-gain	c	
1	50	Self-emp-n...	83311	Bachelors	13	Married-civ...	Exec-mana...	Husband	White	Male	0	0	
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cl...	Not-in-family	White	Male	0	0	
3	53	Private	234721	11th	7	Married-civ...	Handlers-cl...	Husband	Black	Male	0	0	
4	28	Private	338409	Bachelors	13	Married-civ...	Prof-specialty	Wife	Black	Female	0	0	
5	37	Private	284582	Masters	14	Married-civ...	Exec-mana...	Wife	White	Female	0	0	
6	49	Private	160187	9th	5	Married-sp...	Other-service	Not-in-family	Black	Female	0	0	
7	52	Self-emp-n...	209642	HS-grad	9	Married-civ...	Exec-mana...	Husband	White	Male	0	0	
8	31	Private	45781	Masters	14	Never-marri...	Prof-specialty	Not-in-family	White	Female	14084	0	
9	42	Private	159449	Bachelors	13	Married-civ...	Exec-mana...	Husband	White	Male	5178	0	
10	37	Private	280464	Some-college	10	Married-civ...	Exec-mana...	Husband	Black	Male	0	0	
11	30	State-gov	141297	Bachelors	13	Married-civ...	Prof-specialty	Husband	Asian-Pac-L...	Male	0	0	
12	23	Private	122272	Bachelors	13	Never-marri...	Adm-clerical	Own-child	White	Female	0	0	
13	32	Private	205019	Assoc-acdm	12	Never-marri...	Sales	Not-in-family	Black	Male	0	0	
14	40	Private	121772	Assoc-voc	11	Married-civ...	Craft-repair	Husband	Asian-Pac-L...	Male	0	0	
15	34	Private	245487	7th-8th	4	Married-civ...	Transport-...	Husband	Amer-India...	Male	0	0	
16	25	Self-emp-n...	176756	HS-grad	9	Never-marri...	Farming-fis...	Own-child	White	Male	0	0	
17	32	Private	186824	HS-grad	9	Never-marri...	Machine-op...	Unmarried	White	Male	0	0	
18	38	Private	28887	11th	7	Married-civ...	Sales	Husband	White	Male	0	0	
19	43	Self-emp-n...	292175	Masters	14	Divorced	Exec-mana...	Unmarried	White	Female	0	0	

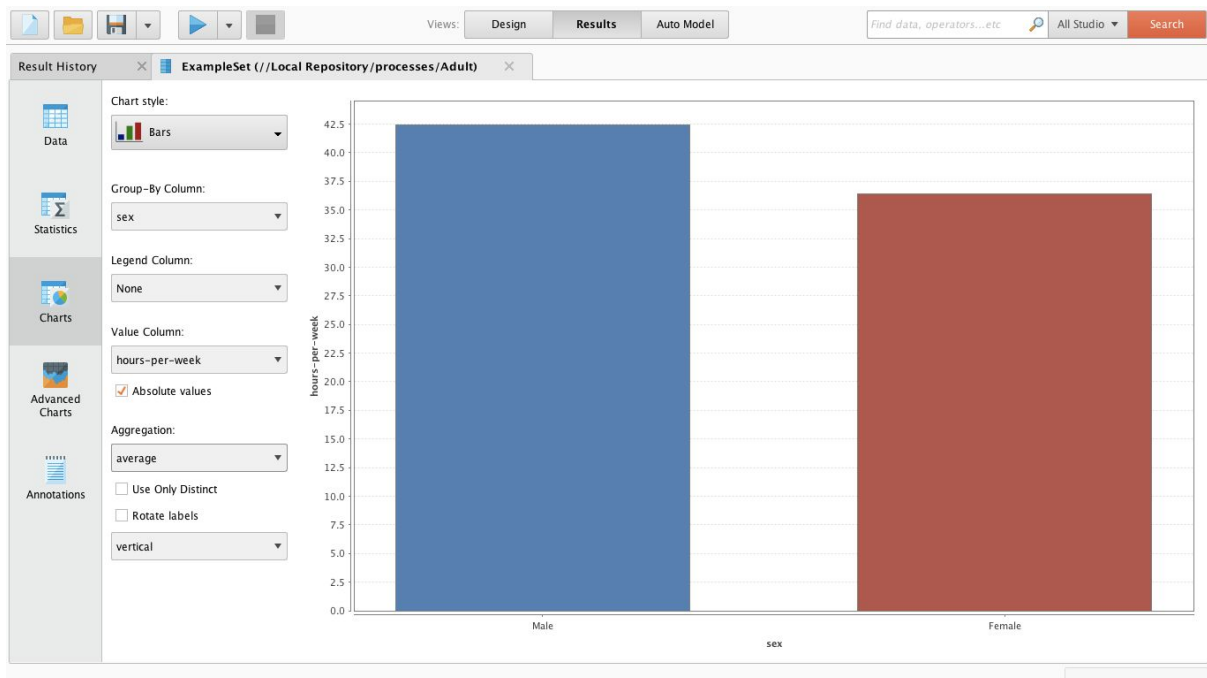
Views: Design Results Auto Model						Find data, operators...etc		All Studio	Search
Result History ExampleSet (//Local Repository/processes/Adult)									
Data	Name	Type	Missing	Statistics	Filter (15 / 15 attributes): Search for Attribute: [v]				
Statistics	occupation	Polynomial	0	Armed-Forces (9)	Prot-specialty (4140)	Prot-specialty (4140), Craft-repair (4099), ...[13 more]			
Charts	relationship	Polynomial	0	Least Other-relative (981)	Most Husband (13193)	Values Husband (13193), Not-in-family (8304), ...[4 more]			
Advanced Charts	race	Polynomial	0	Least Other (271)	Most White (27815)	Values White (27815), Black (3124), ...[3 more]			
Annotations	sex	Polynomial	0	Least Female (10771)	Most Male (21789)	Values Male (21789), Female (10771)			
	capital-gain	Integer	0	Min 0	Max 99999	Average 1077.615			
	capital-loss:	Integer	0	Min 0	Max 4356	Average 87.307			
	hours-per-week	Integer	0	Min 1	Max 99	Average 40.437			
	native-country	Polynomial	0	Least Holand-Netherlands (1)	Most United-States (29169)	Values United-States (29169), Mexico (643), ...[40 more]			
	Perceptions	Polynomial	0	Least >50K (7841)	Most <=50K (24719)	Values <=50K (24719), >50K (7841)			
Showing attributes 1 - 15									
Examples: 32,560 Special Attributes: 0 Regular Attributes: 15									

GRÁFICA 1

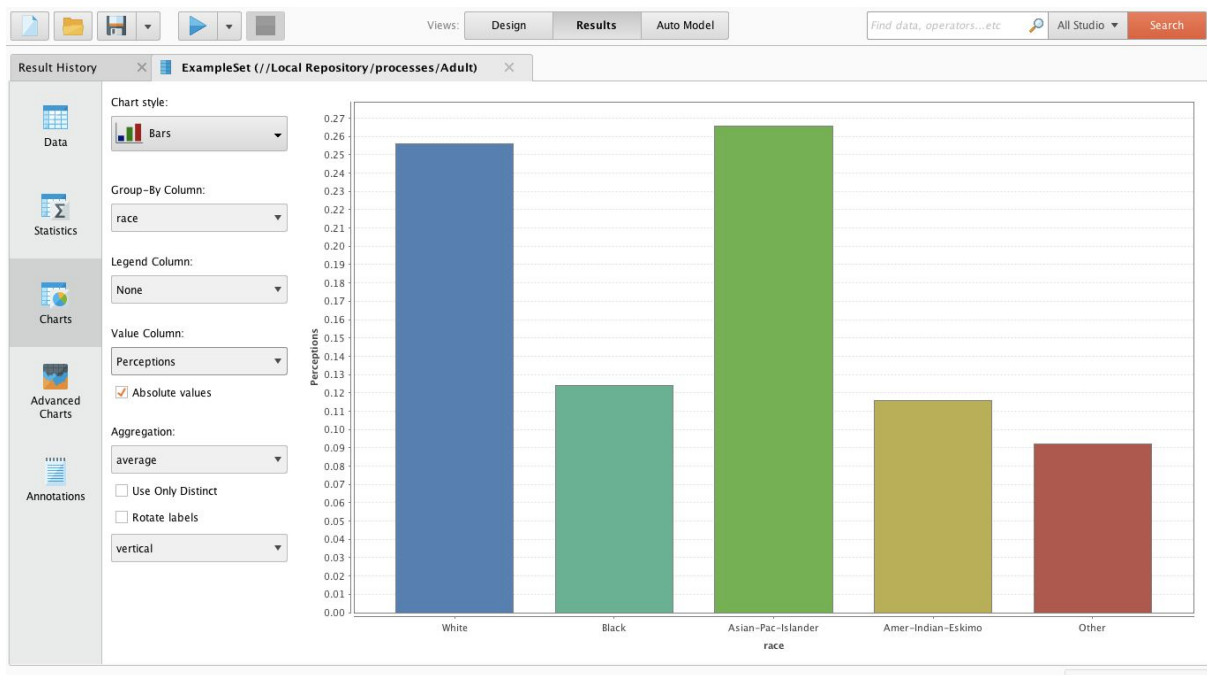


Número de individuos dependiendo el tipo de color de acuerdo a la ganancia capital. Los datos anomalas serían los que se indican como categoría other.

GRÁFICA 2

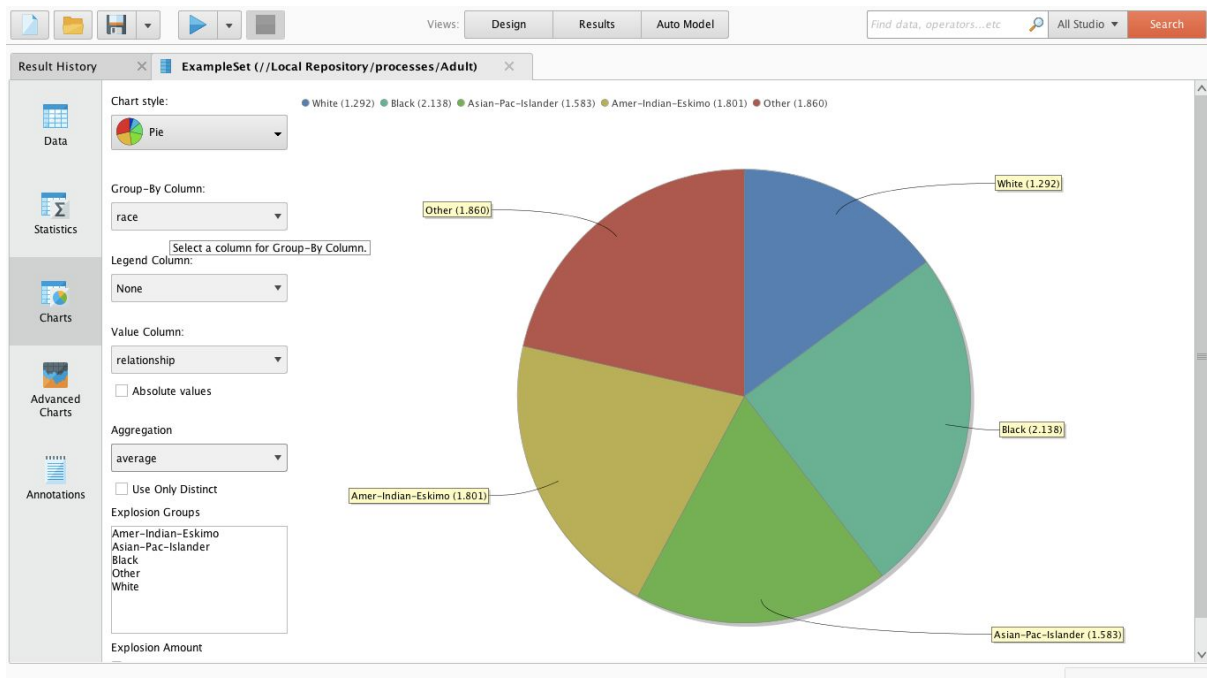


Promedio de horas que se trabajan por género por semana. No hay datos inconsistentes.
GRÁFICA 3



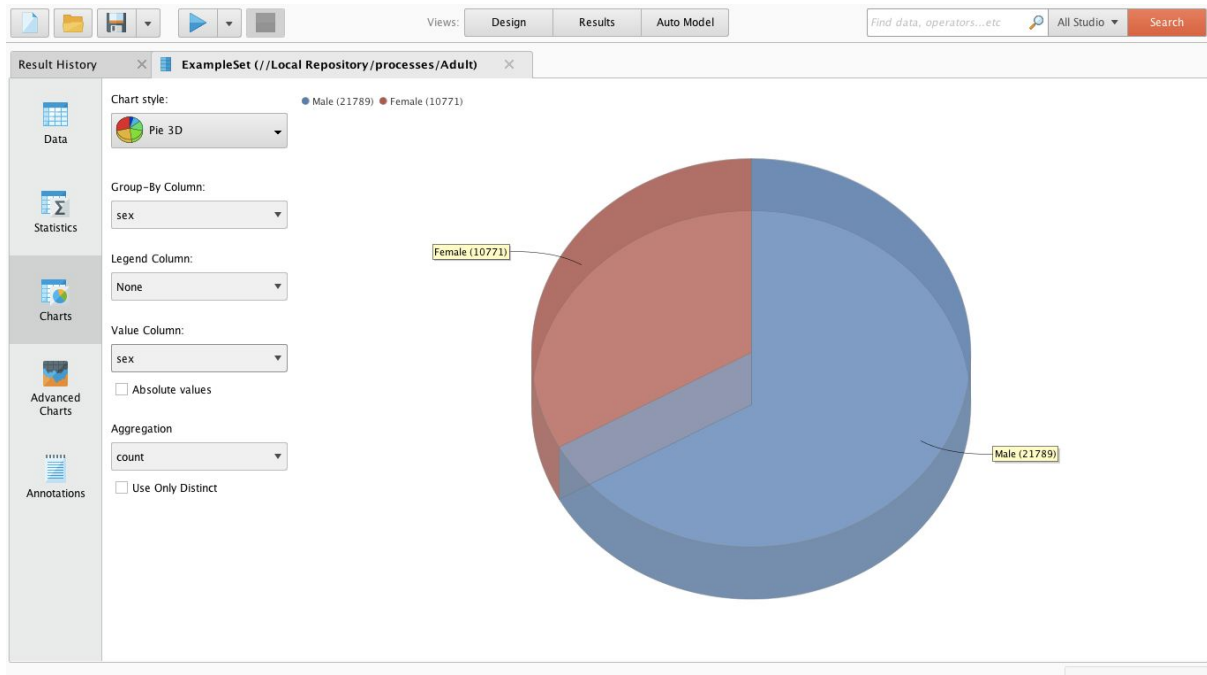
Promedio de percepciones de acuerdo a los individuos por tipo de piel. Como dato anómala podemos distinguir las percepciones de estos, ya que son datos menores que uno.

GRÁFICA 4



Promedio de las personas con relación de acuerdo al tipo de piel. Los datos anormales serían aquellos en los que recae en la categoría *other*.

GRÁFICA 5



Hace la comparación entre la cantidad de individuos por género, hombre y mujeres. De esta manera nos damos cuenta de que hay más hombres que mujeres. Por otra parte, nos damos cuenta que no hay datos anómalos.