

NOMBRE: _____

Los archivos necesarios para hacer este test están en el archivo comprimido TEST1.ZIP que hay en el campus virtual. El código se ha de introducir en la plantilla que se facilita. Las respuestas concretas las debéis incluir en esta hoja y entregarlo al finalizar la prueba. También tendréis que enviar la plantilla completada utilizando el enlace que hay en el Campus Virtual antes de la hora límite.

Ejercicio 1.

El fichero ATP2015.CSV tiene datos de los partidos de semifinales y finales de torneos de tenis de la ATP entre enero y septiembre de 2015. Los datos están separados por punto y coma (;) y los campos, que aparecen en la primera línea, son los siguientes:

VARIABLE	DESCRIPCIÓN
Torneo	Nombre del torneo
Superficie	Tipo de superficie (Clay=Tierra batida, Hard=Pista Rápida, Grass=Hierba)
Nivel	Modalidad del torneo (1=ATP World Tour, 2=Masters 1000, 3=Grand Slam)
Fecha	Fecha del partido
Ganador	
Gmano	Mano que usa el ganador (1=Zurdo, 2=Diestro, 3=Ambas)
Galtura	Altura del ganador
Gnacionalidad	Nacionalidad del ganador
Gedad	Edad del ganador
GRanking	Ranking ATP del ganador
Perdedor	
Pmano	Mano que usa el perdedor (1=Zurdo, 2=Diestro, 3=Ambas)
Paltura	Altura del perdedor
Pnacionalidad	Nacionalidad del perdedor
Pedad	Edad del perdedor
PRanking	Ranking ATP del perdedor
Ronda	Ronda del torneo (F=Final, SF=Semifinal)

- a) Haz la importación del archivo ATP2015 creando un archivo con el nombre ATP2015.SAS7BDAT en un directorio del disco duro. Lista los registros leídos.

```
libname d 'd:\test1';  
proc import file='d:\atp2015.csv' out=d.atp2015  
    dbms="csv"  
    replace;  
    getnames=yes;  
    delimiter=";";  
run;  
proc contents data=d.atp2015;  
run;  
proc print 1 data=d.atp2015;  
run;
```

- b) Calcula los siguientes estadísticos para las variables indicadas. Has de controlar la posible presencia de datos faltantes.

	Media	Desviación Estándar
Gedad	28.85	3.44
Galtura	186.94	7.93
Pedad	28.04	3.50
Paltura	188.57	8.63

```

proc means data=d.atp2015;
    var Gedad Galtura Pedad Paltura;
run;
* Hay datos faltantes en las variables de altura;
data d.atp2015;
    set d.atp2015;
    if Galtura=-1 then Galtura=.;
    if Paltura=-1 then Paltura=.;
run;
proc means data=d.atp2015 maxdec=2;
    var Gedad Galtura Pedad Paltura;
run;

```

- c) Selecciona los partidos donde el ranking ATP del perdedor es mejor que el del ganador. Haz un listado en que aparezca el *torneo*, la *ronda*, el *ganador*, el *ranking del ganador*, el *perdedor* y el *ranking del perdedor*. El listado no ha de tener número de observación y al final ha de indicar cuantos registros aparecen listados.

¿Cuantos partidos cumplen esa condición? **RESPUESTA:** _____51_____

```

*Creando un fichero temporal con la selección;
data atp2 (where=(Granking>Pranking));
    set d.atp2015;
run;
proc print data=atp2 noobs n;
    var torneo ronda Ganador Granking Perdedor Pranking;
run;

*Una alternativa sin crear el fichero temporal;
proc print data=d.atp2015 (where=(Granking>Pranking)) noobs n;
    var torneo ronda Ganador Granking Perdedor Pranking;
run;

```

- d) Crea una base de datos temporal FINATP que contenga únicamente las finales de los torneos. Con este fichero se quiere hacer un listado en un fichero pdf, llamado FINALESATP.PDF:
- El listado ha de contener el *torneo*, la *fecha*, el *nivel* del torneo, el *ganador*, la *mano que usa el ganador*, el *perdedor* y la *mano que usa el perdedor*.
 - Los campos *nivel*, *mano del ganador* y *mano del perdedor* han de tener las etiquetas que aparecen en el descriptor del fichero.
 - La fecha ha de tener el año sólo con dos dígitos.
 - En el listado no ha de aparecer la fecha del listado ni el nombre del procedimiento de SAS que se usa y el título ha de ser 'Finales ATP 2015'.

```

options nodate;
ods noproctitle;
ods pdf file='d:\Finalesatp2015.pdf';
title 'Finales ATP 2015';
proc format;
    value nivelf 1='ATP World Tour'
                2='Master 1000'
                3='Grand Slam';
    value manof 1='Zurdo'
                2='Diestro'
                3='Ambas';
run;
data finatp (where=(Ronda='F'));
    set d.atp2015;
    format nivel nivelf. Gmano manof. Pmano manof.;
run;
proc print data=finatp;
    var torneo fecha nivel ganador Gmano perdedor Pmano;
    format fecha ddmmyy8.;
run;
ods pdf close;

```

- e) Con el fichero creado con las finales de los torneos realiza una descriptiva de la nacionalidad del ganador.

¿Qué nacionalidad/es ha/n ganado más torneos? **RESPUESTA:** ___ESP/SRB/SUI con 8_____

Solo teniendo en cuenta las finales de los torneos del Masters 1000 y Grand Slam, haz una tabla de frecuencias con los ganadores de estos torneos.

¿Qué tenista/s ha/n ganado más torneos y cuántos han sido? **RESPUESTA:** Nova Djokovic con 6

```
proc freq data=finatp;
    table gnacionalidad;
run;

proc freq data=finatp (where=(nivel=2 | nivel=3));
    table ganador;
run;
```

Ejercicio 2.

El fichero CLASE.TXT tiene datos de los alumnos de una asignatura. Los campos son los siguientes:

VARIABLE	DESCRIPCIÓN
nom	Nombre
ape1	Primer Apellido
ape2	Segundo Apellido
fecnac	Fecha de nacimiento (dd/mm/yy)
nota	Nota del examen
grup	Grupo de laboratorio

- a) Crea el fichero CLASE.SAS7BDAT en un directorio del disco duro, a partir de la lectura del fichero de texto. Añade etiquetas de variable con las descripciones anteriores. Completa el formato de fecha indicando el año con 4 dígitos. Haz un listado del fichero con las etiquetas como cabeceras de campo.

```
Libname d 'd:\ ';
data d.clase;
    infile 'd:\clase.txt';
    input nom$ 1-12 ape1$ 13-24 ape2$ 25-36 @37 fecnac :ddmmyy8. nota @52 grup$;
    format fecnac :ddmmyy10.;
    label nom='Nombre'
           ape1='Primer Apellido'
           ape2='Segundo Apellido'
           fecnac='Fecha de Nacimiento'
           nota= 'Nota del Examen'
           grup= 'Grupo';
run;
*Alternativamente input nom$ 12. ape1$ 12. ape2$ 12. @37 fecnac :ddmmyy8. nota @52
grup$;
proc print 1 data=d.clase;
run;
```

- b) Haz una tabla de frecuencias de las *notas* obtenidas teniendo en cuenta la categorización clásica:

Nota	ETIQUETA
Menos de 5	SUSPENSO
[5,7)	APROBADO
[7,9)	NOTABLE
9 ó más	SOBRESALIENTE

¿Qué porcentaje de SUSPENSOS hay? **RESPUESTA:** _48.42%_____

```
proc format;
    value notaf low-<5="Suspenso"
                5-<7="Aprobado"
                7-<9="Notable"
                9-high="Sobresaliente";
run;
proc freq data=d.clase;
    table nota;
    format nota notaf.;
run;
```

- c) Haz un listado de clase ordenado por *grupo*, *primer apellido*, *segundo apellido* y *nombre*, donde además de estos campos aparezca la *nota* del examen.

```
proc sort data=d.clase out=clase_ord;
    by grup apel ape2 nom;
run;

proc print 1 data=clase_ord;
    var grup apel ape2 nom nota;
run;
```

- d) Crea dos bases de datos temporales, una que se llame CLASEA y que contenga toda la información de los alumnos del grupo A, y otra base que se denomine CLASEB con los alumnos del grupo B. Para cada una de estas bases, obtén la descriptiva de la *nota* del examen.

¿Qué grupo tiene *nota* media más alta? **RESPUESTA:** __El A con 5.47__

```
data clasea (where=(grup="A"));
    set d.clase;
run;
proc means data=clasea maxdec=2;
    var nota;
run;

data claseb (where=(grup="B"));
    set d.clase;
run;
proc means data=claseb maxdec=2;
    var nota;
run;
```

- e) Lista los *nombres*, *apellidos* y las *notas* del grupo A, ordenando los alumnos de la mejor nota a la peor.

```
proc sort data=clasea out=clasea_ord;
    by descending nota;
run;
proc print 1 data=clasea_ord;
    var nom apel ape2 nota;
run;
```

¿Qué alumno ha conseguido la nota más alta? **RESPUESTA:** __Jordi Mota Gómez con 9.8__