

ACADEMIA DE STUDII ECONOMICE DIN BUCUREȘTI FACULTATEA DE CIBERNETICĂ, STATISTICĂ ȘI INFORMATICĂ ECONOMICĂ



한류사랑축제

The World On
arirang

Prof. univ. dr. Adina Ileana UȚĂ
Studenta: Giol Adriana
Grupa 1078, Seria C

București, 2020

Conținutul proiectului (partea a III-a – Programare SAS)

<i>1. Crearea Unui Set De Date Sas Din Fișiere Externe</i>	<i>3</i>
<i>2. Crearea Și Folosirea De Formate Definite De Utilizator</i>	<i>12</i>
<i>3. Generare De Grafice</i>	<i>14</i>
<i>4. Proceduri Statistice</i>	<i>16</i>
<i>5. Proceduri Pentru Raportare</i>	<i>21</i>
<i>6. Interogări Sql</i>	<i>23</i>
<i>7. Utilizare De Funcții Sas Și Procesare Condițională A Datelor</i>	<i>25</i>
<i>8. Lucrul Cu Masive (Array) Și Procesare Iterativă A Datelor</i>	<i>27</i>
<i>9. Subseturi De Date Cu Procesare Condițională A Datelor</i>	<i>29</i>
<i>Bibliografie</i>	<i>30</i>

PARTEA a III-a – PROGRAMARE SAS

1. CREAREA UNUI SET DE DATE SAS DIN FIȘIERE EXTERNE

• Descrierea problemei

Managerul televiziunii ArirangTV dorește să importe în SAS Studio fișierul CSV, "Raiting", care conține date despre filmul "Parasite" și să salveze aceste date permanent pentru modificări ulterioare.

De asemenea, în mod asemănător, acesta dorește și importul următoarelor fișiere CSV: "Filme", "Actori", "Rand Actori", "Cheltuieli Publicitare".

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ Denumirea continentului.
- ✓ *MedieRaiting* – Media raitingului obținut.
- ✓ *NrVoturi* – Numărul de persoane care au acordat un raiting filmului.
- ✓ *BileteVândute* - Numărul de bilete vândute pe fiecare continent.
- ✓ *MediePrețBilet* – Prețul mediu al unui bilet pe fiecare dintre continente.

	A	B	C	D	E	F
1	ID	Continent	MedieRaiting	NrVoturi	BileteVandute	MediePretBilet
2	1	America de Nord	6.8	9256366	15623545	15.74
3	2	America de Sud	7.5	1000452	6780225	10.99
4	3	Australia	7.1	701478	4521520	20.26
5	4	Asia	8.6	1566421	10636665	12.65
6	5	Africa	7.1	502365	1125556	6.44
7	6	Europa	8.2	998256	5899119	10.1

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC IMPORT
- ✓ Metoda de calcul folosită:

Am încărcat fișierul CSV Raiting în folderul Proiect și cu opțiunea Properties din meniul contextual al fișierului importat i-am aflat calea. În meniul Snippets, în folderul Data am găsit codul pentru importul de fișiere CSV. Am modificat acest cod pentru a se potrivi cu cerințele avute în vedere. Setul de date temporar numit Raiting s-a creat în meniul Libraries, în folderul Work, acesta fiind șters odată ce sesiunea curentă este închisă.

Pentru a crea un set de date permanent vom crea o librărie nouă numită LibPro, având ca și cale-rădăcină folderul Proiect. Am rulat din nou codul și astfel am creat un set de date permanent numit Raiting.

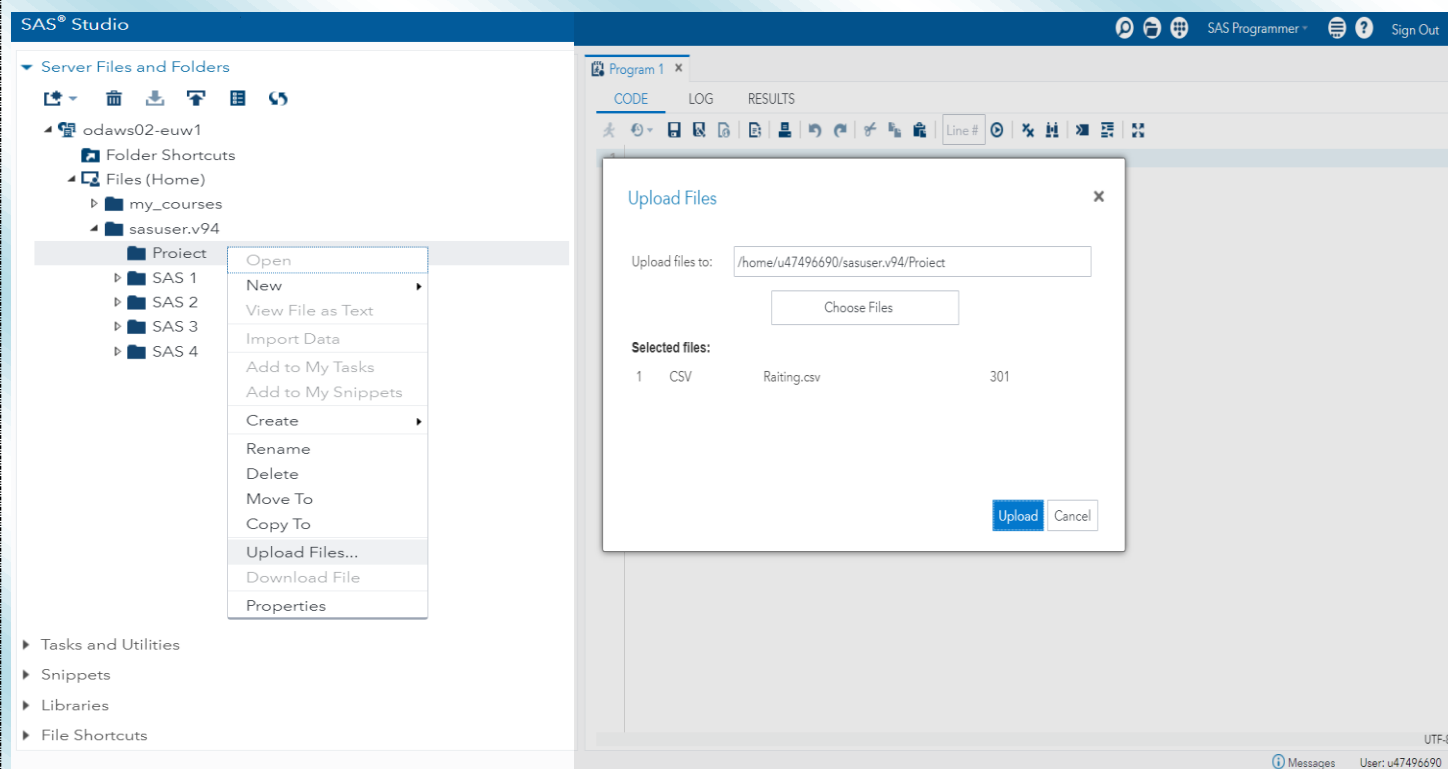
În mod asemănător fișierului CSV Raiting am importat încă 4 fișiere CSV pentru prelucrări ulterioare:

- Fișierul Filme: Conține filmele câștigătoare de la Korean Drama Awards, indiferent de categoria de premiu câștigat, și detalii despre aceste filme.
- Fișierul Actori: Conține numele și detaliile personale ale celor mai buni actori, colaboratori cu ArirangTV.

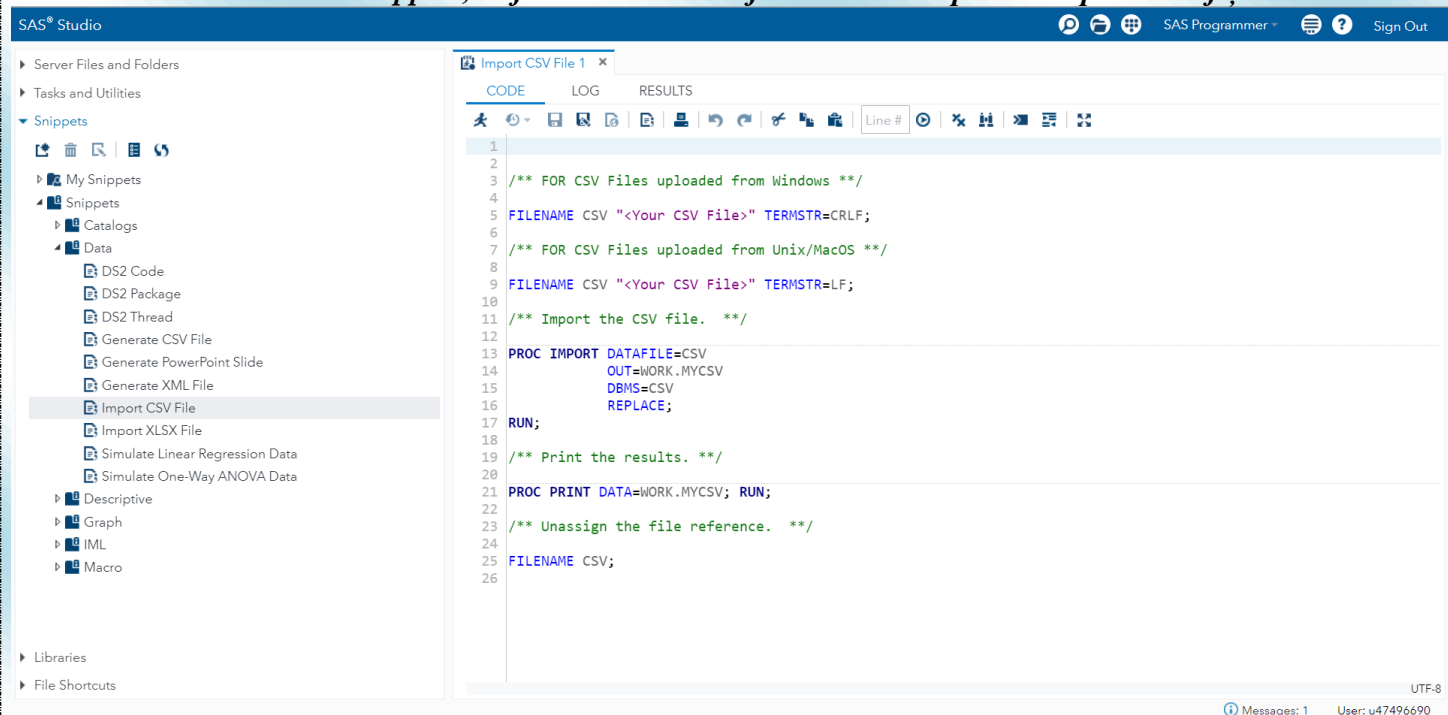
- **Fișierul Rând Actori:** Reprezintă o corespondență între actorii și filmele jucate de aceștia, astfel spus, o corespondență între fișierul Filme și fișierul Actori, conținând și salariul obținut de actori pentru acel film.
- **Fișierul Cheltuieli Publicitare:** Conține cheltuielile înregistrate pentru filme în primele 6 luni și numărul de bilete vândute pentru acestea.

✓ **Metoda de calcul folosită pe pași:**

- **PAS 1: Încărcarea fișierului CSV în Folderul Proiect și aflarea căi-rădăcină.**



- **PAS 2: În Meniul Snippets, în folderul Data se află codul sursă pentru importul de fișiere CSV.**



➤ **PAS 3: Modificăm acest cod pentru cerințele de care avem nevoie.**

```
1. Import fisiere CSV.sas x
CODE LOG RESULTS OUTPUT DATA
1 /*1. IMPORTUL FISIERELOR CSV*/
2
3 /*1. Fisier RAITING*/
4 /* Pentru fisiere CSV incarcate din Windows */
5 FILENAME CSV "/home/u47496690/sasuser.v94/Proiect/Raiting.csv" TERMSTR=CRLF;
6
7 /*Import fisier CSV - Raiting */
8 PROC IMPORT DATAFILE=CSV OUT=LIBPRO.Raiting DBMS=CSV REPLACE;
9 RUN;
10
11 /** Afisare rezultate **/
12 Title'Tabel RAITING ~ Detaliile filmului "Parasite";
13 PROC PRINT DATA=LIBPRO.Raiting;
14 RUN;
15
16 /** Anularea referintei fisierului **/
17 FILENAME CSV;
18
```

➤ **PAS 4: După rularea codului s-a creat setul de date temporar numit Raiting în meniul Libraries, în folderul Work, acesta fiind șters odată ce sesiunea curentă este închisă.**

SAS® Studio

Server Files and Folders

Tasks and Utilities

Snippets

Libraries

My Libraries

MAPS

MAPSGFK

MAPSSAS

SASDATA

SASHELP

SASUSER

STPSAMP

WEBWORK

WORK

RAITING

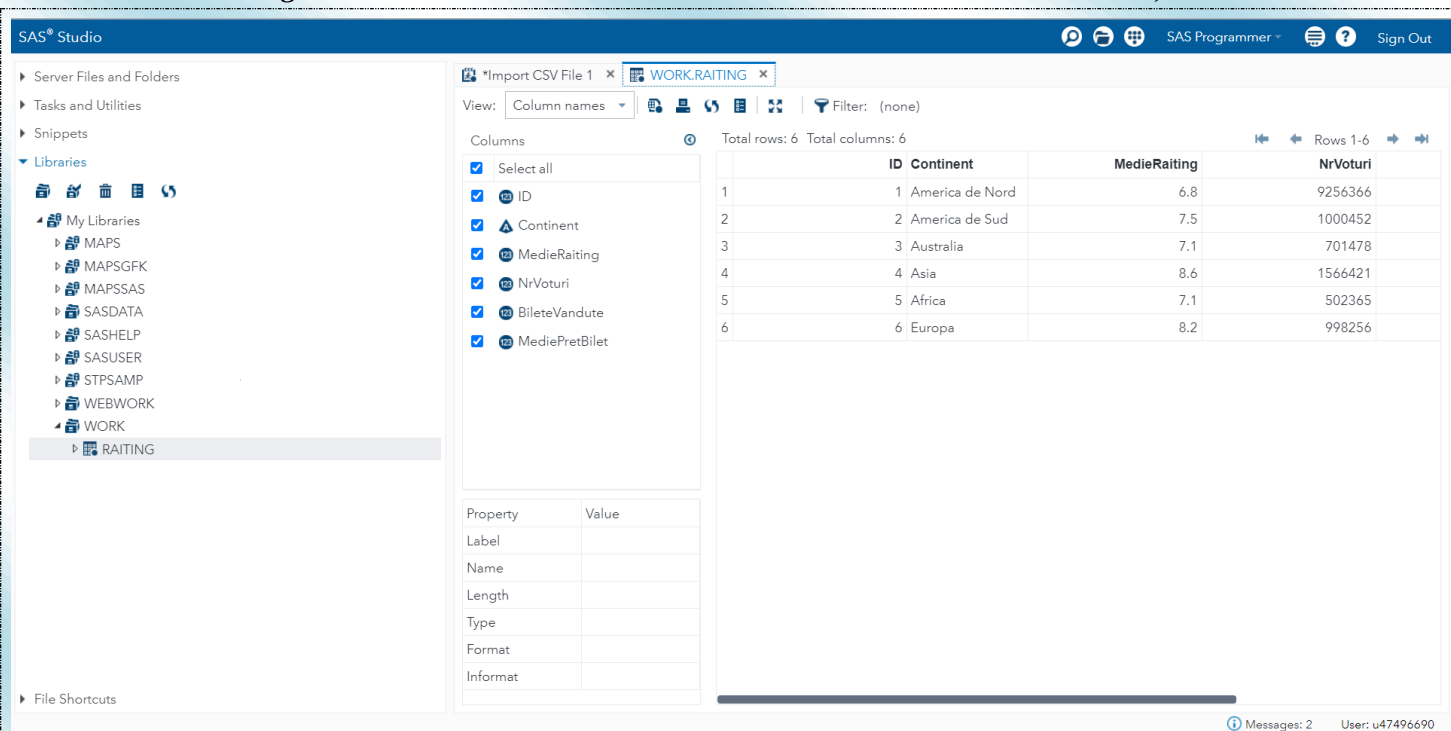
Import CSV File 1 x WORK.RAITING x

CODE LOG RESULTS OUTPUT DATA

Table of Contents

Obs	ID	Continent	MedieRaiting	NrVoturi	BileteVandute	MediePretBilet
1	1	America de Nord	6.8	9256366	15623545	15.74
2	2	America de Sud	7.5	1000452	6780225	10.99
3	3	Australia	7.1	701478	4521520	20.26
4	4	Asia	8.6	1566421	10636665	12.65
5	5	Africa	7.1	502365	1125556	6.44
6	6	Europa	8.2	998256	5899119	10.1

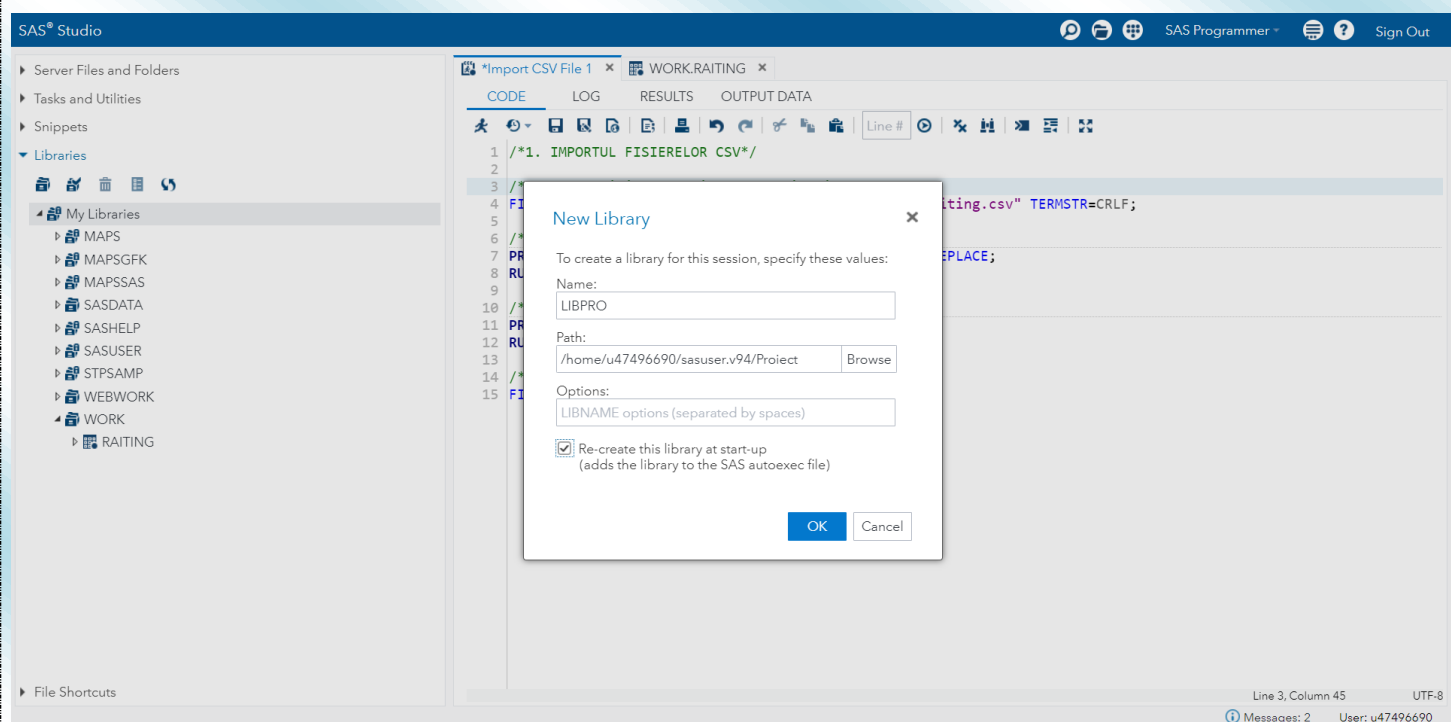
Messages: 2 User: u47496690



The screenshot shows the SAS Studio interface. On the left, the 'Libraries' pane is visible, showing a tree structure with 'My Libraries' and 'WORK' (containing 'RAITING'). The main window displays the 'Import CSV File' dialog for 'WORK.RAITING'. The 'Columns' list includes 'ID', 'Continent', 'MedieRaiting', 'NrVoturi', 'BileteVandute', and 'MediePretBilet'. The 'View' is set to 'Column names'. The resulting data table is shown with 6 rows and 6 columns.

	ID	Continent	MedieRaiting	NrVoturi	
1	1	America de Nord	6.8	9256366	
2	2	America de Sud	7.5	1000452	
3	3	Australia	7.1	701478	
4	4	Asia	8.6	1566421	
5	5	Africa	7.1	502365	
6	6	Europa	8.2	998256	

➤ **PAS 5: Pentru a crea un set de date permanent vom crea o librărie nouă numită LIBPRO, având ca și cale-rădăcină folderul Proiect.**



The screenshot shows the SAS Studio interface with the 'New Library' dialog box open. The dialog box prompts the user to create a library for this session, specifying values for Name, Path, and Options. The 'Name' field is set to 'LIBPRO', the 'Path' is '/home/u47496690/sasuser.v94/Proiect', and the 'Options' field is 'LIBNAME options (separated by spaces)'. The 'Re-create this library at start-up' checkbox is checked.

➤ **PAS 6: Am rulat din nou codul și astfel am creat un set de date permanent numit Raiting, care se află în biblioteca proprie LIBPRO.**

SAS® Studio

Server Files and Folders

Tasks and Utilities

Snippets

Libraries

My Libraries

LIBPRO

RAITING

MAPS

MAPSGFK

MAPSSAS

SASDATA

SASHELP

SASUSER

STPSAMP

WEBWORK

WORK

File Shortcuts

*1. Import fisiere CSV.sas

CODE LOG RESULTS OUTPUT DATA

```

1 /*1. IMPORTUL FISIERELOR CSV*/
2
3 /** Pentru fisiere CSV incarcate din Windows **/
4 FILENAME CSV "/home/u47496690/sasuser.v94/Proiect/Raiting.csv" TERMSTR=CRLF;
5
6 /** Import fisier CSV - Raiting **/
7 PROC IMPORT DATAFILE=CSV OUT=LIBPRO.Raiting DBMS=CSV REPLACE;
8 RUN;
9
10 /** Afisare rezultate **/
11 PROC PRINT DATA=LIBPRO.Raiting;
12 RUN;
13
14 /** Anularea referintei fisierului **/
15 FILENAME CSV;
    
```

/home/u47496690/sasuser.v94/Proiect/1. Import fisiere CSV.sas

Line 15, Column 14 UTF-8

Messages: 2 User: u47496690

SAS® Studio

Server Files and Folders

Tasks and Utilities

Snippets

Libraries

My Libraries

LIBPRO

RAITING

MAPS

MAPSGFK

MAPSSAS

SASDATA

SASHELP

SASUSER

STPSAMP

WEBWORK

WORK

File Shortcuts

*1. Import fisiere CSV.sas

CODE LOG RESULTS OUTPUT DATA

Table of Contents

Obs	ID	Continent	MedieRaiting	NrVoturi	BileteVandute	MediePretBilet
1	1	America de Nord	6.8	9256366	15623545	15.74
2	2	America de Sud	7.5	1000452	6780225	10.99
3	3	Australia	7.1	701478	4521520	20.26
4	4	Asia	8.6	1566421	10636665	12.65
5	5	Africa	7.1	502365	1125556	6.44
6	6	Europa	8.2	998256	5899119	10.1

Messages: 2 User: u47496690

➤ **PAS 7: De asemenea, se poate observa și în folderul Proiect că a apărut un fișier numit *raiting.sas7bdat*, extensia *.sas7bdat* fiind formatul principal utilizat pentru stocarea seturilor de date SAS. Extensia de fișier "sas7bdat" a fost introdusă începând cu versiunea 7 de SAS.**

The screenshot shows the SAS Studio interface. On the left, the 'Server Files and Folders' pane displays the project structure, including the file 'raiting.sas7bdat'. The main window shows the 'Columns' pane with a list of variables: ID, Continent, MedieRaiting, NrVoturi, BileteVandute, and MediePretBilet. The 'View' dropdown is set to 'Column names'. The 'Total rows: 6' and 'Total columns: 6' are displayed. The data table shows the following rows:

	ID	Continent	MedieRaiting	NrVoturi
1	1	America de Nord	6.8	9256366
2	2	America de Sud	7.5	1000452
3	3	Australia	7.1	701478
4	4	Asia	8.6	1566421
5	5	Africa	7.1	502365
6	6	Europa	8.2	998256

Codul utilizat

```
/*1. IMPORTUL FISIERELOR CSV*/

/*1. Fisier RAITING*/
/* Pentru fisiere CSV incarcate din Windows */
FILENAME CSV "/home/u47496690/sasuser.v94/Proiect/Raiting.csv" TERMSTR=CRLF;
/*Import fisier CSV - Raiting */
PROC IMPORT DATAFILE=CSV OUT=LIBPRO.Raiting DBMS=CSV REPLACE;
RUN;
/** Afisare rezultate **/
Title'Tabel RAITING ~ Detaliile filmului "Parasite"';
PROC PRINT DATA=LIBPRO.Raiting;
RUN;
/** Anularea referintei fisierului **/
FILENAME CSV;

/*****/

/*2. Fisier FILME*/
/** Pentru fisiere CSV incarcate din Windows **/
FILENAME CSV "/home/u47496690/sasuser.v94/Proiect/Filme.csv" TERMSTR=CRLF;
/** Import fisier CSV - Filme **/
PROC IMPORT DATAFILE=CSV OUT=LIBPRO.Filme DBMS=CSV REPLACE;
RUN;
/** Afisare rezultate **/
Title'Tabel FILME ~ Filmele produse de ArirangTV';
PROC PRINT DATA=LIBPRO.Filme;
RUN;
/** Anularea referintei fisierului **/
FILENAME CSV;

/*****/
```



```
/*3. Fisier ACTORI*/
/** Pentru fisiere CSV incarcate din Windows **/
FILENAME CSV "/home/u47496690/sasuser.v94/Proiect/Actori.csv" TERMSTR=CRLF;
/** Import fisier CSV - Actori **/
PROC IMPORT DATAFILE=CSV OUT=LIBPRO.Actori DBMS=CSV REPLACE;
RUN;
/** Afisare rezultate **/
Title'Tabel ACTORI ~ Actorii principali din productiile ArirangTV';
PROC PRINT DATA=LIBPRO.Actori;
RUN;
/** Anularea referintei fisierului **/
FILENAME CSV;
/*****/
/*4. Fisier RAND ACTORI*/
/** Pentru fisiere CSV incarcate din Windows **/
FILENAME CSV "/home/u47496690/sasuser.v94/Proiect/RandActori.csv" TERMSTR=CRLF;
/** Import fisier CSV - RandActori **/
PROC IMPORT DATAFILE=CSV OUT=LIBPRO.RandActori DBMS=CSV REPLACE;
RUN;
/** Afisare rezultate **/
Title'Tabel RAND ACTORI ~ Corespondenta dintre actori si filmele in care au jucat';
PROC PRINT DATA=LIBPRO.RandActori;
RUN;
/** Anularea referintei fisierului **/
FILENAME CSV;
/*****/
/*5. Fisier CHELTUIELI PUBLICITARE*/
/** Pentru fisiere CSV incarcate din Windows **/
FILENAME CSV "/home/u47496690/sasuser.v94/Proiect/CheltuieliPublicitare.csv"
TERMSTR=CRLF;
/** Import fisier CSV - Cheltuieli Publicitare **/
PROC IMPORT DATAFILE=CSV OUT=LIBPRO.CheltuieliPublicitare DBMS=CSV REPLACE;
RUN;
/** Afisare rezultate **/
Title'Tabel CHELTUIELI PUBLICITARE ~ Cheltuielile publicitare pentru filmele ArirangTV';
PROC PRINT DATA=LIBPRO.CheltuieliPublicitare;
RUN;
/** Anularea referintei fisierului **/
FILENAME CSV;
```

Rezultatele obtinute



SAS Programmer



Sign Out

1. Import fisiere CSV.sas

CODE

LOG

RESULTS

OUTPUT DATA



Table of Contents

Tabel RAITING ~ Detaliile filmului "Parasite"

Obs	ID	Continent	MedieRaiting	NrVoturi	BileteVandute	MediePretBilet
1	1	America de Nord	6.8	9256366	15623545	15.74
2	2	America de Sud	7.5	1000452	6780225	10.99
3	3	Australia	7.1	701478	4521520	20.26
4	4	Asia	8.6	1566421	10636665	12.65
5	5	Africa	7.1	502365	1125556	6.44
6	6	Europa	8.2	998256	5899119	10.1

Tabel FILME ~ Filmele produse de ArirangTV

Obs	IdFilm	Film	An	Regizor	Buget	Raiting
1	1	A Frozen Flower	2009	Yoo Ha	25000000	7.1
2	2	A Werewolf Boy	2012	Jung Soo Hee	130000000	8.5
3	3	Along with the Gods	2018	Kim Yong Hwa	150000000	8.9
4	4	Assassination	2015	Lee Ki Chul	19000000	6.3
5	5	Midnight Runners	2017	Kim Joo Hwan	140000000	7.8
6	6	What a Man Wants	2018	Lee Byung Hun	135000000	6.8
7	7	Train to Busan	2016	Yeon Sang Ho	135000000	9.1

Tabel ACTORI ~ Actorii principali din productiile ArirangTV

Obs	IdActor	Nume	Gen	DataNasterii	Nationalitate
1	1	Seo Kang Joon	m	12/10/1993	coreeana
2	2	Lee Na Young	f	22/02/1979	coreeana
3	3	Lee Jong Suk	m	14/09/1989	coreeana
4	4	Jung Jessica	f	18/04/1989	americana
5	5	Jung Krystal	f	24/10/1994	americana
6	6	Park Hae Soo	m	21/11/1981	coreeana
7	7	Manoban Lalisa	f	27/03/1997	thailandeza
8	8	Hirai Momo	f	09/11/1996	japoneza
9	9	Zhou Tzu Yu	f	14/06/1999	taiwaneza
10	10	Song Joong Ki	m	19/09/1985	coreeana
11	11	Kim Ji Won	f	19/10/1992	coreeana
12	12	Jang Dong Gun	m	07/03/1972	coreeana
13	13	Kim Sang Joong	m	06/10/1965	coreeana
14	14	Lee Min Ho	m	22/06/1987	coreeana

Tabel RAND ACTORI ~ Corespondenta dintre actori si filmele in care au jucat

Obs	IdFilm	IdActor	Salariu
1	1	1	1000000
2	1	2	15000000
3	2	3	2100000
4	2	4	750000
5	3	5	1200000
6	3	6	1300000
7	4	9	1000000
8	4	7	400000
9	5	8	950000
10	5	9	1000000
11	6	10	800000
12	6	14	750000
13	6	11	5000000
14	6	12	850000
15	7	13	1400000
16	7	14	1100000

Tabel CHELTUIELI PUBLICITARE ~ Cheltuielile publicitare pentru filmele ArirangTV

Obs	Film	Luna	Cheltuieli	BileteVandute
1	A Frozen Flower	1	250000	500000
2	A Frozen Flower	2	350000	450123
3	Train to Busan	2	270000	300000
4	What a Man Wants	2	410000	1000000
5	What a Man Wants	3	500000	998700
6	Train to Busan	3	100000	150000
7	Assassination	3	310000	170000
8	Train to Busan	4	800000	1500000
9	Assassination	4	100000	110000
10	Train to Busan	5	750000	1450000
11	Midnight Runners	6	1000000	2100510
12	A Werewolf Boy	6	800000	1200000
13	Along with the Gods	6	1110000	2000000

- **Interpretare:**

Se poate observa că în folderul Proiect a apărut un fișier numit rating.sas7bdat și acest fișier conține datele cerute de către managerul televiziunii ArirangTV despre filmul "Parasite". În mod asemănător, s-a creat și pentru celealte fișiere CSV importate câte un fișier cu extensia .sas7bdat.

2. CREAREA ȘI FOLOSIREA DE FORMATE DEFINITE DE UTILIZATOR

• Descrierea problemei

Managerul televiziunii ArirangTV dorește să afișeze datele din fișierul *Raiting* folosind următorul format:

- Dacă *raiting-ul* obținut de filmul "Parasite" pe un anumit continent este mai mic decât 6.5, se va afișa mesajul "Scăzut".
- Dacă *raiting-ul* se regăsește între 6.6 și 8, se va afișa "Mediu".
- Dacă *raiting-ul* depășește 8.1, rezultatul va fi "Ridicat".

De asemenea, prețul mediu al biletelor va avea ca monedă dolarul (\$).

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ *Raiting-ul* obținut de filmul "Parasite".
- ✓ Prețul mediu al biletelor.

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC FORMAT
- ✓ Metoda de calcul folosită:

Cu ajutorul procedurii PROC FORMAT am creat un format definit de utilizator numit "raiting" care grupează *raiting-ul* obținut de filmul "Parasite" în 3 categorii de *raiting* "Scăzut", "Mediu", "Ridicat". Acest format l-am aplicat prin rularea blocului *Data*, care salvează aceste modificări în librărie anterior creată, LIBPRO. L-a sfârșit am menționat ca media prețului unui bilet să aibe ca unitate de măsură dolarul.

The screenshot displays the SAS Studio environment. On the left, the 'Server Files and Folders' pane shows a project structure with files like '1. Import fisiere CSV.sas', '2. Formate definite de utilizator.sas', and others. The main editor window is titled '2. Formate definite de utilizator.sas' and contains the following SAS code:

```

/*2. CREAREA SI FOLOSIREA DE FORMATE DEFINITE DE UTILIZATOR*/

/*Afișarea datelor din fișierul Raiting folosind un format personalizat */
PROC FORMAT;
value raiting low-6.5='Scăzut'
              6.6-8 = 'Mediu'
              8.1-high ='Ridicat';

RUN;
TITLE 'Clasificarea Raiting-ului';
PROC PRINT data=libpro.raiting label;
label  MedieRaiting="Raiting mediu"
       NrVoturi ="NumarVoturi"
       BileteVandute ="Numar bilete vandute"
       MediePretBilet ="Pret mediu bilet";
format MedieRaiting raiting. MediePretBilet dollar10.2;

RUN;
  
```

The bottom status bar indicates the file path: /home/u47496690/sasuser.v94/Proiect/2. Formate definite de utilizator.sas, and shows 2 messages and the user u47496690.

- Codul utilizat

```
/*2. CREAREA SI FOLOSIREA DE FORMATE DEFINITE DE UTILIZATOR*/
```

```
/*Afisarea datelor din fisierul Raiting folosind un format personalizat */
```

```
PROC FORMAT;
```

```
value raiting low-6.5='Scazut'
```

```
        6.6-8 = 'Mediu'
```

```
        8.1<-high ="Ridicat";
```

```
RUN;
```

```
TITLE'Clasificarea Raiting-ului';
```

```
PROC PRINT data=libpro.raiting label;
```

```
label  MedieRaiting="Raiting mediu"
```

```
        NrVoturi ="NumarVoturi"
```

```
        BileteVandute ="Numar bilete vandute"
```

```
        MediePretBilet ="Pret mediu bilet";
```

```
format MedieRaiting raiting. MediePretBilet dollar10.2;
```

```
RUN;
```

- Rezultatele obtinute

SAS Studio interface showing the results of the SAS program. The main window displays the output of the PROC PRINT statement, which is a table titled "Clasificarea Raiting-ului". The table contains 6 observations (rows) and 7 columns (headers): Obs, ID, Continent, Raiting mediu, NumarVoturi, Numar bilete vandute, and Pret mediu bilet. The data shows that Asia and Europa have the highest rating (Ridicat), while America de Nord and America de Sud have a medium rating (Mediu).

Obs	ID	Continent	Raiting mediu	NumarVoturi	Numar bilete vandute	Pret mediu bilet
1	1	America de Nord	Mediu	9256366	15623545	\$15.74
2	2	America de Sud	Mediu	1000452	6780225	\$10.99
3	3	Australia	Mediu	701478	4521520	\$20.26
4	4	Asia	Ridicat	1566421	10636665	\$12.65
5	5	Africa	Mediu	502365	1125556	\$6.44
6	6	Europa	Ridicat	998256	5899119	\$10.10

- Interpretare:

Se poate observa că raiting-ul cel mai ridicat s-a înregistrat pe continentele Asia și Europa.

3. GENERARE DE GRAFICE

• Descrierea problemei

Managerul televiziunii ArirangTV dorește să realizeze un grafic pentru a vizualiza numărul de bilete vândute pe fiecare continent pentru filmul "Parasite".

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ Denumirea continentului.
- ✓ Numărul de bilete vândute pe fiecare continent.

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC TEMPLATE, PROC SGRENDER
- ✓ Metoda de calcul folosită:

În meniul Tasks and Utilities, folderul Tasks, subfolderul Graph am ales opțiunea Pie Chart. În fereastra deschisă am ales ca și furnizor de date Libraria LIBPRO.Raiting, iar la roluri am ales ca și Categorie numele continentului, iar ca și variabilă BileteleVândute (numărul de bilete vândute). La rulare s-a creat codul din imagine care a creat graficul de mai jos, conform cerințelor managerului.

The screenshot displays the SAS Studio interface. On the left, the 'Tasks and Utilities' pane shows the 'Graph' folder expanded, with 'Pie Chart' selected. The main workspace is divided into three panes: 'DATA', 'APPEARANCE', and 'INFORMATION'. The 'DATA' pane shows the data source as 'LIBPRO.RAITING' and the role as 'Continent'. The 'APPEARANCE' pane shows the measure as 'BileteVandute' and the statistic as 'Sum (default)'. The 'INFORMATION' pane shows the variable as 'BileteVandute'. On the right, the 'CODE' pane displays the generated SAS code:

```

1  * Task code generated by SAS Studio 3.8
2  *
3  * Generated on '5/1/20, 12:00 AM'
4  * Generated by 'u47496690'
5  * Generated on server 'ODANS01-EUW1.ODA.SAS.COM'
6  * Generated on SAS platform 'Linux X64 3.10.0-1062.9.1.el7'
7  * Generated on SAS version '9.04.01M6P11072018'
8  * Generated on browser 'Mozilla/5.0 (Windows NT 10.0; Win64; x
9  * Generated on web client 'https://odamid-euw1.oda.sas.com/SAS
10 *
11 */
12
13 /* Define Pie template */
14 proc template;
15   define statgraph SASStudio.Pie;
16     begingraph;
17       layout region;
18       piechart category=Continent response=BileteVandute /;
19     endlayout;
20   endgraph;
21 run;
22
23 ods graphics / reset width=6.4in height=4.8in imagemap;
24
25 proc sgrender template=SASStudio.Pie data=LIBPRO.RAITING;
26 run;
27
28 ods graphics / reset;
29

```

• Codul utilizat

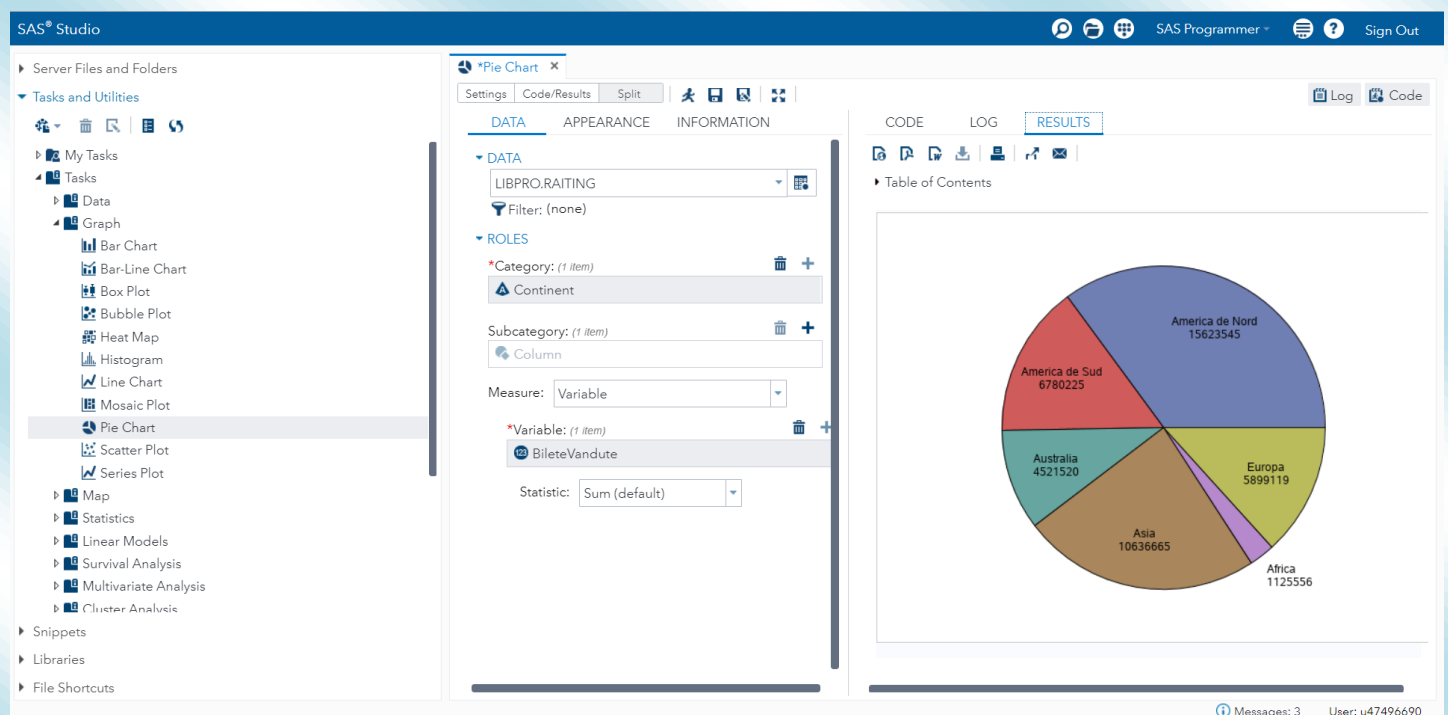
```
/*3.GRAFIC*/
/* Define Pie template */
/* Crearea unui Pie Chart in functie de numarul biletelor vandute pe fiecare continent */
proc template;
    define statgraph SASStudio.Pie;
        begingraph;
            layout region;
            piechart category=Continent response=BileteVandute /;
            endlayout;
        endgraph;
    end;
run;

ods graphics / reset width=6.4in height=4.8in imagemap;

proc sgrender template=SASStudio.Pie data=LIBPRO.RAITING;
run;

ods graphics / reset;
```

• Rezultatele obtinute



• Interpretare:

După cum se poate observa, America de Nord este continentul cu cele mai multe bilete vândute la filmul "Parasite", adunând un număr de 15.623.545 bilete. Continentul cu cel mai puțin interes față de film este Africa, având numai 1.125.556 bilete.

4. PROCEDURI STATISTICE

- Descrierea problemei

Managerul televiziunii ArirangTV dorește o analiză statistică a datelor din fișierul Raiting din punct de vedere al prețului mediu al biletelor de pe fiecare continent.

- Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ *Prețul mediu al biletelor.*
- ✓ *Denumirea continentului.*

- Produs software / procedură / metodă de calcul folosită

- ✓ *Produs software folosit: SAS Studio*
- ✓ *Limbaj: SAS*
- ✓ *Procedură: PROC UNIVARIATE:*

Procedura UNIVARIATE, parte a BASE SAS, este folosită, în principal, pentru examinarea distribuției datelor pentru o singură variabilă, incluzând o evaluare a normalității datelor și identificarea valorilor extreme. Pe lângă calcularea unor indicatori statistici de bază (care se vor regăsi și în procedura MEANS), UNIVARIATE poate include și reprezentări grafice necesare pentru analiza datelor.

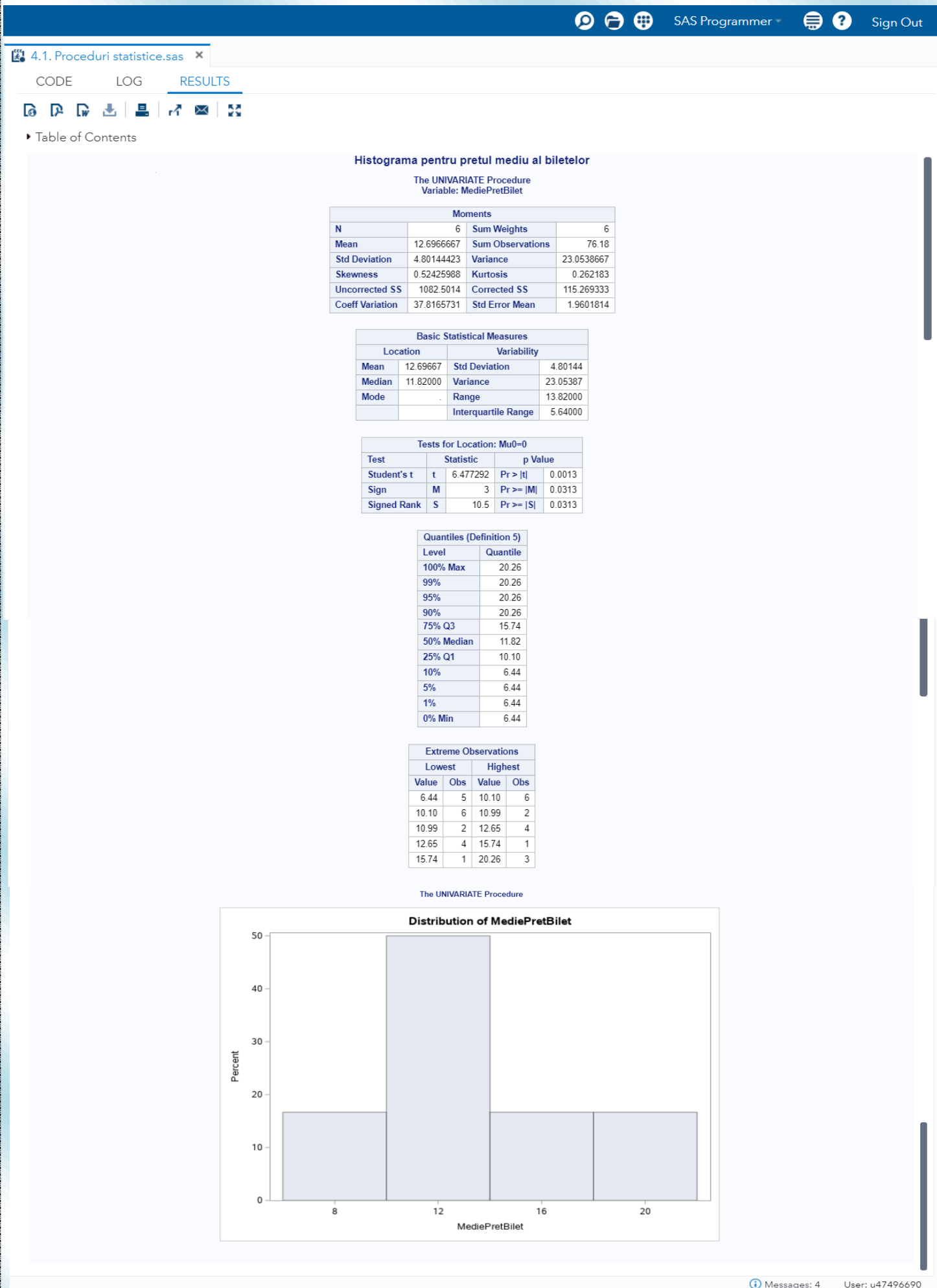
- ✓ *Metoda de calcul folosită:*

- **ANALIZA 1:** *În primă fază, am folosit procedura UNIVARIATE pentru a analiza statistic datele din fișierul Raiting din punct de vedere al prețului mediu al biletelor de pe fiecare continent.*

- Codul utilizat

```
/*4.1. PROCEDURI STATISTICE*/  
/*Histograma pentru pretul mediu al biletelor*/  
PROC UNIVARIATE DATA=LIBPRO.Raiting;  
  VAR MediePretBilet;  
  HISTOGRAM MediePretBilet;  
  Title "Histograma pentru pretul mediu al biletelor";  
RUN;
```


Rezultatele obtinute



• Interpretare:

După cum se poate observa din imaginea de mai sus (Tabel 2: Basic Statistical Measures) media prețului unui bilet la nivel global este de 12.69 dolari, în timp ce mediana este 11.82 dolari, valori foarte apropiate între ele.

În Tabelul 3: Quantiles (Definition 5) se poate observa că la nivelul cuartilei 3 avem o sumă de 15.74 de dolari, iar cuartila 1 valorează 10.10 dolari.

Din graficul de mai sus rezultă că aproximativ 50% din bilete au o valoare apropiată de 12 dolari, în timp ce restul de 50% se împart în mod aproximativ egal între celelalte prețuri posibile.

- **ANALIZA 2:** În continuare, am decis să realizez același tip de analiză statistică. Înainte de a face acest lucru, am creat o nouă tabelă SAS, denumită Vanzari în care, pe lângă coloanele care se regăsesc și în tabelul Raiting, am adăugat două coloane noi:

★ Vanzările totale - Numărul de bilete vândute \times Prețul mediu al unui bilet.

★ Procentul de vot - Numărul de voturi / Numărul de bilete vândute

De asemenea, am afișat rezultatele obținute printr-un format personalizat.

La final, am aplicat procedura UNIVARIATE

• Codul utilizat

```
/*4.2 PROCEDURI STATISTICE*/
/*Crearea unei noi tabele SAS - Vanzari*/
libname LIBPRO '/home/u47496690/sasuser.v94/Proiect';
data LIBPRO.Vanzari;
    set libpro.Raiting;
    VanzariTotale=BileteVandute*MediePretBilet;
    ProcentVot=NrVoturi/BileteVandute;
RUN;

/*Aplicarea unui format predefinit*/
PROC FORMAT;
value raiting low-6.5='Scazut'
              6.6-8 = 'Mediu'
              8.1<-high ="Ridicat";
RUN;

PROC PRINT data=libpro.Vanzari label;
label  MedieRaiting="Raiting mediu"
        NrVoturi ="NumarVoturi"
        BileteVandute ="Numar bilete vandute"
        MediePretBilet ="Pret mediu bilet"
        VanzariTotale = "Vanzari Totale"
        ProcentVot="Procent de vot";
format MedieRaiting raiting. MediePretBilet dollar10.2 ProcentVot percent10.2;
RUN;

/*Aplicarea analizei statistice*/
PROC UNIVARIATE DATA=LIBPRO.Vanzari;
VAR ProcentVot;
HISTOGRAM ProcentVot;
Title"Histograma pentru procentul de vot";
RUN;
```

Rezultatele obtinute

SAS Programmer Sign Out

4.2. Proceduri Statistice.sas

CODE LOG RESULTS OUTPUT DATA

Table of Contents

Obs	ID	Continent	Raiting mediu	NumarVoturi	Numar bilete vandute	Pret mediu billet	Vanzari Totale	Procent de vot
1	1	America de Nord	Mediu	9256366	15623545	\$15.74	245914598.30	59.25%
2	2	America de Sud	Mediu	1000452	6780225	\$10.99	74514672.75	14.76%
3	3	Australia	Mediu	701478	4521520	\$20.26	91605995.20	15.51%
4	4	Asia	Ridicat	1566421	10636665	\$12.65	134553812.25	14.73%
5	5	Africa	Mediu	502365	1125556	\$6.44	7248580.64	44.63%
6	6	Europa	Ridicat	998256	5899119	\$10.10	59581101.90	16.92%

Histograma pentru procentul de vot

The UNIVARIATE Procedure
Variable: ProcentVot

Moments			
N	6	Sum Weights	6
Mean	0.27632875	Sum Observations	1.65797253
Std Deviation	0.19402985	Variance	0.03764758
Skewness	1.1997775	Kurtosis	-0.4602006
Uncorrected SS	0.6463834	Corrected SS	0.18823792
Coeff Variation	70.2170331	Std Error Mean	0.07921236

Basic Statistical Measures			
Location		Variability	
Mean	0.276329	Std Deviation	0.19403
Median	0.162182	Variance	0.03765
Mode	.	Range	0.44520
		Interquartile Range	0.29877

Tests for Location: Mu0=0			
Test	Statistic		p Value
Student's t	t	3.488455	Pr > t 0.0175
Sign	M	3	Pr >= M 0.0313
Signed Rank	S	10.5	Pr >= S 0.0313

Quantiles (Definition 5)	
Level	Quantile
100% Max	0.592463
99%	0.592463
95%	0.592463
90%	0.592463
75% Q3	0.446326
50% Median	0.162182
25% Q1	0.147554
10%	0.147266
5%	0.147266
1%	0.147266
0% Min	0.147266

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.147266	4	0.147554	2
0.147554	2	0.155142	3
0.155142	3	0.169221	6
0.169221	6	0.446326	5
0.446326	5	0.592463	1

Histograma pentru procentul de vot

The UNIVARIATE Procedure

Distribution of ProcentVot

Messages: 7 User: u47496690

- **Interpretare:**

După cum se poate observa din imaginea de mai sus (Tabel 2: Basic Statistical Measures), la nivel global, în medie, numai 27% dintre persoanele care au cumparat bilet să vizioneze filmul "Parasite" și-au exprimat opinia legată de acest film. De asemenea, mediana la nivel global este de 16%, ceea ce înseamnă că pe aproximativ jumătate din continente, procentul de vot a fost mult mai mic.

Graficul susține cele menționate mai sus, sugerând că pe aproximativ 60% din continentele analizate s-au acordat rating-uri în proporție de 0.18%.

5. PROCEDURI PENTRU RAPORTARE

• Descrierea problemei

Managerul televiziunii ArirangTV a investit foarte mulți bani în publicitatea producțiilor ArirangTV de aceea, acesta dorește să vadă dacă eforturile sale au adus beneficii materiale. Prin urmare, dorește un raport prin care să evalueze cheltuielile cu publicitatea în ultimele 6 luni și numărul de bilete vândute, tot în aceeași perioadă de timp.

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ Cheltuielile publicitare.
- ✓ Numărul de bilete vândute.

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC MEANS:

Procedura MEANS are rolul de a crea rapoarte agregate prin generarea de diferiți indicatori statistici pentru variabilele de tip numeric. De asemenea, poate fi folosită pentru a crea seturi de date agregate care pot fi ulterior analizate cu ajutorul altor secțiuni de date sau proceduri.

- ✓ Metoda de calcul folosită:

Cum această procedură permite analiza mai multor variabile, am utilizat-o pentru a crea un raport statistic privind cheltuielile lunare cu publicitatea și rezultatele acestora, concretizate în numărul de bilete vândute.

• Codul utilizat

```
/*5. PROCEDURI PENTRU RAPORTARE*/
libname LIBPRO '/home/u47496690/sasuser.v94/Proiect';
data LIBPRO.CheltuieliPublicitare;
    set LIBPRO.CheltuieliPublicitare;
/*Calcularea mediei cheltuielilor lunare cu publicitatea si a biletelor vandute*/
PROC MEANS DATA=libpro.cheltuielipublicitare;
    BY Luna;
    VAR Cheltuieli BileteVandute;
    TITLE 'Raportul cheltuielilor lunare publicitare si a biletelor vandute';
RUN;
```

• Rezultatele obtinute

SAS Programmer ? Sign Out

5. Proceduri pentru raportare.sas ×

CODE
LOG
RESULTS
OUTPUT DATA

► Table of Contents

Raportul cheltuielilor lunare publicitare si a biletelor vandute

The MEANS Procedure

Luna=1

Variable	N	Mean	Std Dev	Minimum	Maximum
Cheltuieli	1	250000.00	.	250000.00	250000.00
BileteVandute	1	500000.00	.	500000.00	500000.00

Luna=2

Variable	N	Mean	Std Dev	Minimum	Maximum
Cheltuieli	3	343333.33	70237.69	270000.00	410000.00
BileteVandute	3	583374.33	368533.50	300000.00	1000000.00

Luna=3

Variable	N	Mean	Std Dev	Minimum	Maximum
Cheltuieli	3	303333.33	200083.32	100000.00	500000.00
BileteVandute	3	439566.67	484326.92	150000.00	998700.00

Luna=4

Variable	N	Mean	Std Dev	Minimum	Maximum
Cheltuieli	2	450000.00	494974.75	100000.00	800000.00
BileteVandute	2	805000.00	982878.43	110000.00	1500000.00

Luna=5

Variable	N	Mean	Std Dev	Minimum	Maximum
Cheltuieli	1	750000.00	.	750000.00	750000.00
BileteVandute	1	1450000.00	.	1450000.00	1450000.00

Luna=6

Variable	N	Mean	Std Dev	Minimum	Maximum
Cheltuieli	3	970000.00	157162.34	800000.00	1110000.00
BileteVandute	3	1766836.67	493460.66	1200000.00	2100510.00

i Messages: 3 User: u47496690

• Interpretare:

După cum se poate observa, media cheltuielilor lunare cu publicitatea a crescut lunar semnificativ, în timp ce numărul de bilete nu a fost întotdeauna în creștere. Cu toate acestea, în ultima lună, numărul de bilete a atins cote impresionante.

6. INTEROGĂRI SQL

• Descrierea problemei

Managerul televiziunii ArirangTV dorește o interogare SQL care să afișeze salariul și bonusul total primit de fiecare actor pentru filmele jucate, precum și câștigurile totale ale acestora. De asemenea, mai dorște încă o interogare care să afișeze cheltuielile totale cu salariile și bonusurile pentru fiecare film în parte.

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ Salariile actorilor.
- ✓ Raiting-ul obținut de filmul în care a jucat actorul respectiv.

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC SQL
- ✓ Metoda de calcul folosită:

În prima interogare SQL am creat tabela "Salarii Actori" care conține elementele tabeli "Actori" și, în plus față de aceasta, am adăugat și Bonusul obținut de fiecare actor calculat la suma produselor dintre salariile obținute de actorul respectiv și procentul raiting-ului. Am adăugat de asemenea, și coloana Venituri Totale calculată ca sumă dintre salariul inițial și bonusul total câștigat de actor.

În a doua interogare SQL am creat tabela "Cheltuieli Salariale", având ca tabela de baza, tabela "Filme", adăugând la aceasta cheltuielile salariale totale ale fiecarui film, bonusurile totale obținute de actorii din fiecare film și, din suma acestor două coloane a rezultat coloana "Cheltuieli totale".

• Codul utilizat

/*6.1. Interogare SQL*/

/*Interogare SQL care va afisa salariul si bonusul total primit de fiecare actor pentru filmele jucate, precum si castigurile obtinute */

libname LIBPRO '/home/u47496690/sasuser.v94/Proiect';

PROC SQL;

Create table SalariiActori as

Select distinct a.IdActor, Nume, Gen, Nationalitate,
sum(Salariu) as Salariu_total,
sum(Salariu*Raiting/100) as Bonus_total,
sum(Salariu+Salariu*Raiting/100) as Venituri_totale

From LIBPRO.Actori as a, LIBPRO.RandActori as r, LIBPRO.Filme as f

Where a.IdActor=r.IdActor and r.IdFilm=f.IdFilm

Group by a.IdActor;

QUIT;

PROC PRINT DATA=SalariiActori;

RUN;

/*6.2. Interogare SQL*/

/*Interogare SQL care va afisa cheltuielile totale cu salariile si bonusurile pentru fiecare film in parte */

```

libname LIBPRO '/home/u47496690/sasuser.v94/Proiect';
PROC SQL;
CREATE TABLE CheltuieLiSalarii as
Select distinct f.IdFilm,Film,An,Raiting,
                sum(Salariu)as Salarii_totale,
                sum(Salariu*Raiting/100) as Bonusuri_totale,
                sum(Salariu+Salariu*Raiting/100)as CheltuieLi_totale
From LIBPRO.RandActori as r, LIBPRO.Filme as f
Where r.IdFilm=f.IdFilm
Group by f.IdFilm;
QUIT;
PROC PRINT DATA=CheltuieLiSalarii;
RUN;

```

• Rezultatele obtinute

SAS Programmer

*6.1. Interogari SQL.sas

CODE LOG RESULTS OUTPUT DATA

Table of Contents

Obs	IdActor	Nume	Gen	Nationalitate	Salariu_total	Bonus_total	Venituri_totale
1	1	Seo Kang Joon	m	coreeana	1000000	71000	1071000
2	2	Lee Na Young	f	coreeana	15000000	1065000	16065000
3	3	Lee Jong Suk	m	coreeana	2100000	178500	2278500
4	4	Jung Jessica	f	americana	750000	63750	813750
5	5	Jung Krystal	f	americana	1200000	106800	1306800
6	6	Park Hae Soo	m	coreeana	1300000	115700	1415700
7	7	Manoban Lalisa	f	thailandeza	400000	25200	425200
8	8	Hirai Momo	f	japoneza	950000	74100	1024100
9	9	Zhou Tzu Yu	f	taiwaneza	2000000	141000	2141000
10	10	Song Joong Ki	m	coreeana	800000	54400	854400
11	11	Kim Ji Won	f	coreeana	5000000	340000	5340000
12	12	Jang Dong Gun	m	coreeana	850000	57800	907800
13	13	Kim Sang Joong	m	coreeana	1400000	127400	1527400
14	14	Lee Min Ho	m	coreeana	1850000	151100	2001100

6.2. Interogari SQL.sas

CODE LOG RESULTS OUTPUT DATA

Table of Contents

Obs	IdFilm	Film	An	Raiting	Salarii_totale	Bonusuri_totale	CheltuieLi_totale
1	1	A Frozen Flower	2009	7.1	16000000	1136000	17136000
2	2	A Werewolf Boy	2012	8.5	2850000	242250	3092250
3	3	Along with the Gods	2018	8.9	2500000	222500	2722500
4	4	Assassination	2015	6.3	1400000	88200	1488200
5	5	Midnight Runners	2017	7.8	1950000	152100	2102100
6	6	What a Man Wants	2018	6.8	7400000	503200	7903200
7	7	Train to Busan	2016	9.1	2500000	227500	2727500

• Interpretare:

După cum se poate observa, cele mai mari venituri încasate le-a avut actrița Lee Na Young în valoare de \$16 065 000, iar cele mai mari cheltuieli totale avute de televiziunea ArirangTV s-au înregistrat pentru filmul "A Frozen Flower", cheltuielile pentru acest film ridicându-se până la suma de \$17 136 000.

7. UTILIZARE DE FUNCȚII SAS ȘI PROCESARE CONDIȚIONALĂ A DATELOR

• Descrierea problemei

Managerul televiziunii ArirangTV dorește să știe pentru fiecare actor-colaborator cu televiziunea ArirangTV în ce anotimp este născut pentru a pregăti pentru fiecare actor câte o petrecere tematică de ziua lor în concordanță cu festivalurile coreene potrivit lunii și anotimpului respectiv.

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ Data nasterii actorilor.
- ✓ Lunile în care s-au născut actorii care reies din data nașterii.

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC PRINT, IF-ELSE, YRDIF(), TODAY(), MONTH()
- ✓ Metoda de calcul folosită:

Întai, am creat un set de date permanent numit "Vârste" folosindu-mă de setul de date Actori din librărie proprie LIBPRO. Apoi am adăugat încă 3 variabile comparativ cu tabelul original, "Actori" și anume: Vârsta, Luna și Anotimpul.

Am setat anotimpul cu "Iarnă", iar în funcție de luna în care s-a născut actorul (preluată din data nașterii cu funcția MONTH()) se face următoarea distincție:

- dacă luna este mai mare sau egala(GE) cu 3(luna martie) și mai mică (LT) decât 6 (luna Iunie) atunci e "Primăvară.
- dacă luna se încadrează între lunile 6-9 ale anului atunci e "Vară".
- iar dacă se încadrează între lunile 9-12 atunci e "Toamnă".

La final am afișat toate cele 6 variabile ale Actorilor.






• Codul utilizat


```
/*7. UTILIZARE FUNCTII SAS*/

/*Afisarea varstei tuturor actorilor precum si anotimpul in care s-au nascut*/
libname LIBPRO '/home/u47496690/sasuser.v94/Proiect';
DATA LIBPRO.Varste;
SET LIBPRO.Actori;
Varsta=YRDIF(DataNasterii,TODAY(),'Actual');
Luna=MONTH(DataNasterii);
Anotimp='Iarna';
    if Luna ge 3 and Luna lt 6 then Anotimp='Primavara';
    else if Luna ge 6 and Luna lt 9 then Anotimp='Vara';
    else if Luna ge 9 and Luna lt 12 then Anotimp='Toamna';







RUN;
Title "Varstele actorilor";
PROC PRINT DATA=LIBPRO.Varste noobs;
    var IdActor Nume Gen Nationalitate Varsta Anotimp;
RUN;
```

- Rezultatele obtinute**

   SAS Programmer   Sign Out

7.Utilizare Functii SAS.sas 


CODE LOG **RESULTS** OUTPUT DATA

► Table of Contents

Varstele actorilor

IdActor	Nume	Gen	Nationalitate	Varsta	Anotimp
1	Seo Kang Joon	m	coreeana	26.5580	Toamn
2	Lee Na Young	f	coreeana	41.1936	Iarna
3	Lee Jong Suk	m	coreeana	30.6347	Toamn
4	Jung Jessica	f	americana	31.0429	Prima
5	Jung Krystal	f	americana	25.5251	Toamn
6	Park Hae Soo	m	coreeana	38.4484	Toamn
7	Manoban Lalisa	f	thailandeza	23.1032	Prima
8	Hirai Momo	f	japoneza	23.4809	Toamn
9	Zhou Tzu Yu	f	taiwaneza	20.8868	Vara
10	Song Joong Ki	m	coreeana	34.6210	Toamn
11	Kim Ji Won	f	coreeana	27.5383	Toamn
12	Jang Dong Gun	m	coreeana	48.1557	Prima
13	Kim Sang Joong	m	coreeana	54.5744	Toamn
14	Lee Min Ho	m	coreeana	32.8648	Vara

 Messages: 1 User: u47496690

- Interpretare:**

După cum se poate observa, cei mai mulți actori colaboratori cu televiziunea ArirangTV sunt născuți toamna, o singură actriță este născută iarna, 2 actori vara și 3 primăvara.

8. LUCRUL CU MASIVE (ARRAY) ȘI PROCESARE ITERATIVĂ A DATELOR

• Descrierea problemei

Când managerul televiziunii ArriangTV a verificat seturile de date introduse în SAS a observat că în setul de date Actori genul și naționalitatea acestora sunt scrise cu minuscule (LowerCase). Acesta dorește să modifice acest lucru pentru ca genul și naționalitatea actorilor să fie cu majuscule (Upper Case).

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ Genul actorilor.
- ✓ Naționalitatea actorilor.

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC PRINT, ARRAY, DO, UPCASE
- ✓ Metoda de calcul folosită:

Am creat un set de date permanent în librăria proprie LIBPRO, folosindu-mă de setul de date Actori. Apoi, am stabilit numele Array-ului ca fiind "Majuscule Actori", numărul de elemente ce vreau să aparțină din Array adică 2 și anume Gen și Naționalitate, adică elementele de comandă.

După definirea tabloului am adăugat un LOOP DO pentru a se bucla cele 2 elemente și pentru a aplica funcția UpCase la fiecare variabilă.

La final, am adăugat în cod "drop=i" pentru a șterge din tablă "i"-ul rămas în urma buclării.

• Codul utilizat

```
/*8. LUCRUL CU MASIVE -ARRAY*/
```

```
libname LIBPRO '/home/u47496690/sasuser.v94/Proiect';  
proc print data=LIBPRO.Actori;  
run;  
Data LIBPRO.MajusculeActori (drop= i);  
SET LIBPRO.Actori;  
Array MajusculeActori{2}Gen Nationalitate;  
do i=1 to 2;  
MajusculeActori{i}=Upcase(MajusculeActori{i});  
end;  
run;  
proc print data=LIBPRO.MajusculeActori;  
run;
```

• Rezultatele obtinute

SAS Programmer Sign Out

8. Lucrul cu Masive - Array.sas

CODE LOG RESULTS OUTPUT DATA

Table of Contents

Înainte

Obs	IdActor	Nume	Gen	DataNasterii	Nationalitate
1	1	Seo Kang Joon	m	12/10/1993	coreeana
2	2	Lee Na Young	f	22/02/1979	coreeana
3	3	Lee Jong Suk	m	14/09/1989	coreeana
4	4	Jung Jessica	f	18/04/1989	americana
5	5	Jung Krystal	f	24/10/1994	americana
6	6	Park Hae Soo	m	21/11/1981	coreeana
7	7	Manoban Lalisa	f	27/03/1997	thailandeza
8	8	Hirai Momo	f	09/11/1996	japoneza
9	9	Zhou Tzu Yu	f	14/06/1999	taiwaneza
10	10	Song Joong Ki	m	19/09/1985	coreeana
11	11	Kim Ji Won	f	19/10/1992	coreeana
12	12	Jang Dong Gun	m	07/03/1972	coreeana
13	13	Kim Sang Joong	m	06/10/1965	coreeana
14	14	Lee Min Ho	m	22/06/1987	coreeana

După

Obs	IdActor	Nume	Gen	DataNasterii	Nationalitate
1	1	Seo Kang Joon	M	12/10/1993	COREEANA
2	2	Lee Na Young	F	22/02/1979	COREEANA
3	3	Lee Jong Suk	M	14/09/1989	COREEANA
4	4	Jung Jessica	F	18/04/1989	AMERICANA
5	5	Jung Krystal	F	24/10/1994	AMERICANA
6	6	Park Hae Soo	M	21/11/1981	COREEANA
7	7	Manoban Lalisa	F	27/03/1997	THAILANDEZA
8	8	Hirai Momo	F	09/11/1996	JAPONEZA
9	9	Zhou Tzu Yu	F	14/06/1999	TAIWANEZA
10	10	Song Joong Ki	M	19/09/1985	COREEANA
11	11	Kim Ji Won	F	19/10/1992	COREEANA
12	12	Jang Dong Gun	M	07/03/1972	COREEANA
13	13	Kim Sang Joong	M	06/10/1965	COREEANA
14	14	Lee Min Ho	M	22/06/1987	COREEANA

Messages: 1 User: u47496690

• Interpretare:

După cum se poate observa, modificările cerute de către managerul televiziunii s-au concretizat, prin urmare, genul și naționalitatea actorilor au fost modificate în MAJUSCULE.

9. SUBSETURI DE DATE CU PROCESARE CONDIȚIONALĂ A DATELOR

• Descrierea problemei

Managerul televiziunii ArirangTV dorește să știe care dintre actorii colaboratori cu televiziunea ArirangTV sunt născuți în luni de vară, deoarece acesta dorește ca pentru zilele de naștere ale acestora să pregătească torturi de înghețată. Se dorește, de asemenea, să se știe doar actorii coreeni, deoarece cei străini vor avea 5 zile libere pentru a-și sărbători ziua de naștere alături de familiile din stăinătate.

• Informații necesare pentru rezolvare

Pentru realizarea acestei cerințe avem nevoie de următoarele informații :

- ✓ Naționalitatea actorilor.
- ✓ Anotimpul în care s-au născut.

• Produs software / procedură / metodă de calcul folosită

- ✓ Produs software folosit: SAS Studio
- ✓ Limbaj: SAS
- ✓ Procedură: PROC PRINT, IF
- ✓ Metoda de calcul folosită:

Folosind tabela Vârste, creată anterior, am creat un set de date permanent numit "Actori Văratnici" în librerie proprie LIBPRO și am ales Naționalitatea "coreeană" a actorilor și Anotimpul "vară" ca și formă de filtrare a datelor.

• Codul utilizat

```
/*9. Crearea de subseturi*/
/*Am creat un subset de date care afiseaza actorii coreeni nascuti vara*/
libname LIBPRO '/home/u47496690/sasuser.v94/Proiect';
data LIBPRO.ActoriVaratici;
set LIBPRO.Varste;
if Nationalitate eq 'coreeana' and Anotimp eq 'Vara';
title "Actori coreeni nascuti vara care colaboreaza cu ArirangTV";
proc print data=LIBPRO.ActoriVaratici;
run;
```

• Rezultatele obtinute

Obs	IdActor	Nume	Gen	DataNasterii	Nationalitate	Varsta	Luna	Anotimp
1	14	Lee Min Ho	m	22/06/1987	coreeana	32.8648	6	Vara

• Interpretare:

După cum se poate observa, singurul actor coreean care colaborează cu ArirangTV și este născut vara este Lee Min Ho.

Sfârșitul celei de-a treia părți !

BIBLIOGRAFIE**1. Crearea unui set de date SAS din fișiere externe:**

- Seminar 1 Programare SAS
- Import fișier CSV: <https://www.youtube.com/watch?v=oWBOEy9wSjo>
- Creare set de date permanent: <https://www.youtube.com/watch?v=qqI8iuhgWVc>

2. Crearea și folosirea de formate definite de utilizator:

- Seminar 1 Programare SAS
- <https://video.sas.com/detail/video/5334371788001/creating-user-defined-formats-with-the-format-procedure-in-sas%C2%AE-viya%E2%84%A2>

3. Generare de grafice:

- Seminar 4 Programare SAS
- <https://www.youtube.com/watch?v=WfI8ja-sIds>
- <https://video.sas.com/detail/video/4573023403001/creating-a-bar-chart-using-sas-studio>

4. Folosirea de proceduri statistice:

- Seminar 4 Programare SAS
- <https://video.sas.com/detail/video/3306906230001/summary-statistics-using-sas-studio>
- https://media.online.uga.edu/media/PROC+UNIVARIATE+-+Part+1+%28Example+1%29.mp4/1_cjuqmbhb

5. Proceduri pentru raportare:

- Seminar 4 Programare SAS

6. Interogări SQL:

- Seminar 3 Programare SAS
- https://www.youtube.com/watch?v=hX02sbsmb_w
- https://www.youtube.com/watch?v=hX02sbsmb_w&t=215s
- <https://data-flair.training/blogs/sas-sql/>

7. Utilizare de funcții SAS și procesare condițională a datelor

- Seminar 2 Programare SAS

8. Lucrul cu masive (Array) și procesare iterativă a datelor

- Seminar 2 și 3 Programare SAS
- <https://www.sascrunch.com/sas-arrays.html>
- <https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/242-30.pdf>
- <https://data-flair.training/blogs/sas-array/>
- <https://blogs.sas.com/content/sgf/2019/11/05/how-to-make-the-most-of-arrays-with-sas-software-and-streamline-your-programming/>

9. Subseturi de date cu procesare condițională a datelor

- Seminar 2 Programare SAS

Link-Uri Utile:

- **An Introduction to SAS – University Edition De Ron Cody**
<https://books.google.ro/books?id=hM-OCgAAQBAJ&pg=SA6-PA15&lpg=SA6-PA15&dq=sas+studio+iterative+and+conditioner+processing&source=bl&ots=A0DqteswmI&sig=ACfU3U3D1LwkgLPDfa65KDP-YyDn1WKZ8g&hl=ro&sa=X&ved=2ahUKEwjzmfZfpAhWSiqQKHTrzDkUQ6AEwA3oEC-AgQAQ#v=onepage&q&f=false>

- **SAS Studio 5.1: User's Guide, 2018, SAS Institute Inc., Cary, NC, USA**
<https://documentation.sas.com/api/docsets/webeditorug/5.1/content/webeditorug.pdf?locale=en>