

Aisha Muhammad Nawaz L200921

PySpark Lab 4 - Movies (Using SparkSQL) 8A BSCS MMD

12th March 2024

Question 1 (35 marks) You are provided with the dataset "Movies.csv" that contains information about 1600 movies with properties such as year, length, main actor and actress, director and popularity. Your task is to read the file in SPARK RDD and efficiently perform the following queries on the dataset using Spark (Scala). You must use the spark functionality of persisting data in memory and partitioning. Try to write an efficient query.

```
In [ ]: # #Running on CoLab
!pip install pyspark
!pip install -U -q PyDrive
!apt install openjdk-8-jdk-headless -qq
import os
os.environ['JAVA_HOME'] = '/usr/lib/jvm/java-8-openjdk-amd64'
```

Collecting pyspark
 Downloading pyspark-3.5.1.tar.gz (317.0 MB)
 317.0/317.0 MB 2.4 MB/s eta 0:00:00

Preparing metadata (setup.py) ... done
 Requirement already satisfied: py4j==0.10.9.7 in /usr/local/lib/python3.10/dist-packages (from pyspark) (0.10.9.7)
 Building wheels for collected packages: pyspark
 Building wheel for pyspark (setup.py) ... done
 Created wheel for pyspark: filename=pyspark-3.5.1-py2.py3-none-any.whl size=317488493 sha256=409a092c4e59dde6e063b13b5ad0c62d4ece5393e0506b538f00efa2e2617924
 Stored in directory: /root/.cache/pip/wheels/80/1d/60/2c256ed38ddc2fdd93be545214a63e02fbd8d74fb0b7f3a6
 Successfully built pyspark
 Installing collected packages: pyspark
 Successfully installed pyspark-3.5.1
 The following additional packages will be installed:
 libxtst6 openjdk-8-jre-headless
 Suggested packages:
 openjdk-8-demo openjdk-8-source libnss-mdns fonts-dejavu-extra fonts-nanum fonts-ipafont-gothic
 fonts-ipafont-mincho fonts-wqy-microhei fonts-wqy-zenhei fonts-indic
 The following NEW packages will be installed:
 libxtst6 openjdk-8-jdk-headless openjdk-8-jre-headless
 0 upgraded, 3 newly installed, 0 to remove and 38 not upgraded.
 Need to get 39.7 MB of archives.
 After this operation, 144 MB of additional disk space will be used.
 Selecting previously unselected package libxtst6:amd64.
 (Reading database ... 121752 files and directories currently installed.)
 Preparing to unpack .../libxtst6_2%3a1.2.3-1build4_amd64.deb ...
 Unpacking libxtst6:amd64 (2:1.2.3-1build4) ...
 Selecting previously unselected package openjdk-8-jre-headless:amd64.
 Preparing to unpack .../openjdk-8-jre-headless_8u392-ga-1~22.04_amd64.deb ...
 Unpacking openjdk-8-jre-headless:amd64 (8u392-ga-1~22.04) ...
 Selecting previously unselected package openjdk-8-jdk-headless:amd64.
 Preparing to unpack .../openjdk-8-jdk-headless_8u392-ga-1~22.04_amd64.deb ...
 Unpacking openjdk-8-jdk-headless:amd64 (8u392-ga-1~22.04) ...
 Setting up libxtst6:amd64 (2:1.2.3-1build4) ...
 Setting up openjdk-8-jre-headless:amd64 (8u392-ga-1~22.04) ...
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/jre/bin/orbd to provide /usr/bin/orbd (orbd) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/jre/bin/servertool to provide /usr/bin/servertool (servertool) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/jre/bin/tnameserv to provide /usr/bin/tnameserv (tnameserv) in auto mode
 Setting up openjdk-8-jdk-headless:amd64 (8u392-ga-1~22.04) ...
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/clhsdb to provide /usr/bin/clhsdb (clhsdb) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/extcheck to provide /usr/bin/extcheck (extcheck) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/hsdb to provide /usr/bin/hsdb (hsdb) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/idlj to provide /usr/bin/idlj (idlj) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/javah to provide /usr/bin/javah (javah) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jhat to provide /usr/bin/jhat (jhat) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jsadebugd to provide /usr/bin/jsadebugd (jsadebugd) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/native2ascii to provide /usr/bin/native2ascii (native2ascii) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/schemagen to provide /usr/bin/schemagen (schemagen) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/wsgen to provide /usr/bin/wsgen (wsgen) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/wsimport to provide /usr/bin/wsimport (wsimport) in auto mode
 update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/xjc to provide /usr/bin/xjc (xjc) in auto mode
 Processing triggers for libc-bin (2.35-0ubuntu3.4) ...
 /sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_0.so.3 is not a symbolic link

 /sbin/ldconfig.real: /usr/local/lib/libtbb.so.12 is not a symbolic link

 /sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_5.so.3 is not a symbolic link

 /sbin/ldconfig.real: /usr/local/lib/libtbbbind.so.3 is not a symbolic link

 /sbin/ldconfig.real: /usr/local/lib/libtbbmalloc_proxy.so.2 is not a symbolic link

 /sbin/ldconfig.real: /usr/local/lib/libtbbmalloc.so.2 is not a symbolic link

```
In [ ]: # Import the Libraries we will need
import pyspark
from pyspark.sql import *
from pyspark.sql.functions import *
from pyspark import SparkContext, SparkConf

# Create Spark session and ContextRun PySpark.
# create the session
conf = SparkConf().set("spark.ui.port", "4050")
# create the context
sc = pyspark.SparkContext(conf=conf)
spark = SparkSession.builder.appName("DataFrame").config('spark.ui.port', '4050').getOrCreate()
spark
```

Out[]: **SparkSession - in-memory**
SparkContext

[Spark UI \(http://eeb26ae31149:4050\)](http://eeb26ae31149:4050)

Version

v3.5.1

Master

local[*]

AppName

pyspark-shell

```
In [ ]: df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
# Show the DataFrame
df.show(truncate=False)
```

Year	Length	Title	Genre	Actor	Actress	Director	Popularity	Awards	*Image,,,,
1990	111	Tie Me Up! Tie Me Down!	Comedy	Banderas, Antonio	Abril, Victoria	Almodóvar, Pedro	68	No	NicholasCage.pn
1991	113	High Heels	Comedy	Bosé, Miguel	Abril, Victoria	Almodóvar, Pedro	68	No	NicholasCage.pn
1983	104	Dead Zone, The	Horror	Walken, Christopher	Adams, Brooke	Cronenberg, David	79	No	NicholasCage.png,
1979	122	Cuba	Action	Connery, Sean	Adams, Brooke	Lester, Richard	6	No	seanConnery.png,,
1978	94	Days of Heaven	Drama	Gere, Richard	Adams, Brooke	Malick, Terrence	14	No	NicholasCage.pn
1983	140	Octopussy	Action	Moore, Roger	Adams, Maud	Glen, John	68	No	NicholasCage.pn
1984	101	Target Eagle	Action	Connors, Chuck	Adams, Maud	Loma, José Antonio de la	14	No	NicholasCage.pn
1989	99	American Angels: Baptism of Blood, The	Drama	Bergen, Robert D.	Adams, Trudy	Sebastian, Beverly	28	No	NicholasCage.png,
1985	104	Subway	Drama	Lambert, Christopher	Adjani, Isabelle	Besson, Luc	6	No	NicholasCage.pn
1990	149	Camille Claudel	Drama	Depardieu, Gérard	Adjani, Isabelle	Nuytten, Bruno	32	No	NicholasCage.pn
1982	188	Fanny and Alexander	Drama	Ahlstedt, Börje	Adolphson, Kristina	Bergman, Ingmar	81	Yes	Bergman.png,,
1982	117	Tragedy of a Ridiculous Man	Drama	Tognazzi, Ugo	Aimee, Anouk	Bertolucci, Bernardo	17	No	NicholasCage.pn
1966	103	A Man & a Woman	Drama	Trintignant, Jean-Louis	Aimee, Anouk	Lelouch, Claude	46	Yes	NicholasCage.pn
1986	112	A Man & a Woman: Twenty Years Later	Drama	Trintignant, Jean-Louis	Aimee, Anouk	Lelouch, Claude	49	No	NicholasCage.pn
1966	103	Un Hombre y una Mujer	Drama	Trintignant, Jean-Louis	Aimee, Anouk	Lelouch, Claude	6	Yes	NicholasCage.pn
1985	112	Official Story, The	Drama	Alterio, Hector	Aleandro, Norma	Puenzo, Luiz	39	Yes	NicholasCage.png,
1976	150	Lindbergh Kidnapping Case, The	Drama	Hopkins, Anthony	Alexander, Denise	Kulik, Buzz	51	No	AnthonyHopkins.pn
1929	84	Blackmail	Mystery	Longden, John	Algood, Sara	Hitchcock, Alfred	2	No	alfredHitchcock.p
1963	109	Donovan's Reef	Comedy	Wayne, John	Allen, Elizabeth	Ford, John	62	No	johnWayne.png,,

only showing top 20 rows

```
In [ ]: # OR
peopledf =spark.read.csv("Movies.csv", sep= ";", inferSchema=True, header=True)
peopledf.show()
```

Year	Length	Title	Genre	Actor	Actress	Director	Popularity	Awards	*Image,,,,,
1990	111	Tie Me Up! Tie Me...	Comedy	Banderas, Antonio	Abril, Victoria	Almodóvar, Pedro	68	No	NicholasCage.png,,
1991	113	High Heels	Comedy	Bosch, Miguel	Abril, Victoria	Almodóvar, Pedro	68	No	NicholasCage.png,,
1983	104	Dead Zone, The	Horror	Walken, Christopher	Adams, Brooke	Cronenberg, David	79	No	NicholasCage.png,,
1979	122	Cuba	Action	Connery, Sean	Adams, Brooke	Lester, Richard	6	No	seanConnery.png,,
1978	94	Days of Heaven	Drama	Gere, Richard	Adams, Brooke	Malick, Terrence	14	No	NicholasCage.png,,
1983	140	Octopussy	Action	Moore, Roger	Adams, Maud	Glen, John	68	No	NicholasCage.png,,
1984	101	Target Eagle	Action	Connors, Chuck	Adams, Maud	Loma, Jos	14	No	NicholasCage.png,,
1989	99	American Angels: ...	Drama	Bergen, Robert D.	Adams, Trudy	Sebastian, Beverly	28	No	NicholasCage.png,,
1985	104	Subway	Drama	Lambert, Christopher	Adjani, Isabelle	Besson, Luc	6	No	NicholasCage.png,,
1990	149	Camille Claudel	Drama	Depardieu, Gérard	Adjani, Isabelle	Nuytten, Bruno	32	No	NicholasCage.png,,
1982	188	Fanny and Alexander	Drama	Ahlstedt, Börje	Adolphson, Kristina	Bergman, Ingmar	81	Yes	Bergman.png,,
1982	117	Tragedy of a Ridi...	Drama	Tognazzi, Ugo	Aimee, Anouk	Bertolucci, Bernardo	17	No	NicholasCage.png,,
1966	103	A Man & a Woman	Drama	Trintignant, Jean...	Aimee, Anouk	Lelouch, Claude	46	Yes	NicholasCage.png,,
1986	112	A Man & a Woman: ...	Drama	Trintignant, Jean...	Aimee, Anouk	Lelouch, Claude	49	No	NicholasCage.png,,
1966	103	Un Hombre y una M...	Drama	Trintignant, Jean...	Aimee, Anouk	Lelouch, Claude	6	Yes	NicholasCage.png,,
1985	112	Official Story, The	Drama	Alterio, Hector	Aleandro, Norma	Puenzo, Luiz	39	Yes	NicholasCage.png,,
1976	150	Lindbergh Kidnapp...	Drama	Hopkins, Anthony	Alexander, Denise	Kulik, Buzz	51	No	AnthonyHopkins.png,,
1929	84	Blackmail	Mystery	Longden, John	Algood, Sara	Hitchcock, Alfred	2	No	alfredHitchcock.p...
1963	109	Donovan's Reef	Comedy	Wayne, John	Allen, Elizabeth	Ford, John	62	No	johnWayne.png,,

only showing top 20 rows

```
In [ ]: # 1. Find the total number of movies of each genre. Consider only the movies released after 1960 and have a Length greater than 90 minutes.
# df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
df.filter((df.Year>1960) & (df.Length>90)).groupBy('Genre').count().show()
```

Genre	count
Romance	1
Adventure	4
NULL	1
Drama	434
War	15
Fantasy	1
Mystery	63
Music	21
Science Fiction	26
Horror	26
Western	37
Comedy	242
Action	124
Westerns	5

```
In [ ]: # 2. For each actress, find the movie in which she acted. You must print the names of the movies and the year of the movie.
df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')

result_df = df.groupBy('Actress').agg(
    collect_list('Title').alias('Movies'),
    collect_list('Year').alias('Years')
)

# Show the results
result_df.show(truncate=False)
```

[illegible]

https://htmtopdf.herokuapp.com/ipynbviewer/temp/19373545a88913b6c3bbee1fd377bf2f/SQL_Lab4v2_Pyspark_Movies_L200921.html?t=17102325... 6/19

-----+
only showing top 20 rows

```
In [ ]: # 3. Find the 25 most popular movies released during 1980-1990
df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
df.filter((df.Year>=1980) & (df.Year<=1990)).sort(col('Popularity').desc()).limit(25).collect()

Out[ ]: [Row(Year='1986', Length='103', Title='Fire with Fire', Genre='Drama', Actor='Sheffer, Craig', Actress='Madsen, Virginia', Director='Gibbins, Duncan', Popularity
='9', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1989', Length='102', Title='Dead-Bang', Genre='Action', Actor='Johnson, Don', Actress='Miller, Penelope Ann', Director='Frankenheimer, John', Popularity
='9', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1987', Length='104', Title='Lionheart', Genre='Action', Actor='Stoltz, Eric', Actress='Barrymore, Deborah', Director='Schaffner, Franklin J.', Popularity
='9', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1989', Length='104', Title='Tango & Cash', Genre='Action', Actor='Stallone, Sylvester', Actress='Hatcher, Teri', Director='Konchalovsky, Andrei', Popular
ity='9', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1989', Length='83', Title='Masque of the Red Death', Genre='Horror', Actor='MacNee, Patrick', Actress='Hoak, Clare', Director='Brand, Larry', Popularity
='9', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1988', Length='96', Title='Glitz', Genre='Mystery', Actor='Smits, Jimmy', Actress='Post, Markie', Director=None, Popularity='9', Awards='No', *Image,,,,
='NicholasCage.png,,'),
Row(Year='1988', Length='90', Title='Naked Gun: From the Files of Police Squad!', The', Genre='Comedy', Actor='Nielsen, Leslie', Actress='Presley, Priscilla', Direc
tor='Zucker, David', Popularity='9', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1985', Length='124', Title='Plenty', Genre='Drama', Actor='Dance, Charles', Actress='Streep, Meryl', Director='Schepisi, Fred', Popularity='9', Awards='N
o', *Image,,,,='merylStreep.png,,'),
Row(Year='1985', Length='116', Title='Back to the Future', Genre='Comedy', Actor='Fox, Michael J.', Actress='Thompson, Lea', Director='Zemeckis, Robert', Popularit
y='9', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1981', Length='94', Title='Deadline', Genre='Mystery', Actor='Newman, Barry', Actress=None, Director='Nicholson, Arch', Popularity='9', Awards='No', *Ima
ge,,,,='paulNewman.png,,'),
Row(Year='1989', Length=None, Title='Death Valley Days, Deadly Decision', Genre='Western', Actor='Caan, James', Actress=None, Director=None, Popularity='9', Awards
='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1985', Length='55', Title='Gonzo Presents Muppet Weird Stuff', Genre='Comedy', Actor='Cleese, John', Actress='Andrews, Julie', Director=None, Popularity
='88', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1990', Length='105', Title='Guilty by Suspicion', Genre='Drama', Actor='De Niro, Robert', Actress='Bening, Annette', Director='Winkler, Irwin', Populart
y='88', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1988', Length=None, Title='Five Corners', Genre='Drama', Actor='Robbins, Tim', Actress='Foster, Jodie', Director=None, Popularity='88', Awards='No', *Ima
ge,,,,='NicholasCage.png,,'),
Row(Year='1989', Length='86', Title='Let It Ride', Genre='Comedy', Actor='Dreyfuss, Richard', Actress='Garr, Teri', Director='Pytko, Joe', Popularity='88', Awards
='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1989', Length='90', Title='New Year's Day', Genre='Comedy', Actor='Jaglom, Henry', Actress='Jakobsen, Maggie', Director='Jaglom, Henry', Popularity='88',
Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1980', Length=None, Title='Happy Birthday to Me', Genre='Horror', Actor='Ford, Glenn', Actress='Anderson, Melissa Sue', Director='Thompson, J. Lee', Popu
larity='88', Awards='No', *Image,,,,='glennFord.png,,'),
Row(Year='1989', Length='88', Title='Final Notice', Genre='Mystery', Actor='Gerard, Gil', Actress='Anderson, Melody', Director='Stern, Steven Hilliard', Popularity
='88', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1990', Length='95', Title='Dangerous Pursuit', Genre='Mystery', Actor='Harrison, Gregory', Actress='Powers, Alexandra', Director='Stern, Sandor', Popular
ity='88', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1986', Length='104', Title='Best of Times, The', Genre='Comedy', Actor='Williams, Robin', Actress='Reed, Pamela', Director='Spottiswoode, Roger', Popular
ity='88', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1982', Length='128', Title='Ballad of Narayama, The', Genre='Drama', Actor='Ogata, Ken', Actress='Sakamoto, Sumiko', Director='Imamura, Shohei', Populari
ty='88', Awards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1986', Length='90', Title='Head Office', Genre='Comedy', Actor='Reinhold, Judge', Actress='Seymour, Jane', Director='Finkleman, Ken', Popularity='88', Aw
ards='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1985', Length='161', Title='Out of Africa', Genre='Drama', Actor='Redford, Robert', Actress='Streep, Meryl', Director='Pollack, Sydney', Popularity='88',
Awards='Yes', *Image,,,,='merylStreep.png,,'),
Row(Year='1990', Length=None, Title='Blood in, Blood Out', Genre='Drama', Actor='Penn, Sean', Actress=None, Director='Olmos, Edward James', Popularity='88', Awards
='No', *Image,,,,='NicholasCage.png,,'),
Row(Year='1989', Length='107', Title='Star Trek V: The Final Frontier', Genre='Action', Actor='Shatner, William', Actress='Nichols, Nichelle', Director='Shatner, W
illiam ', Popularity='87', Awards='No', *Image,,,,='NicholasCage.png,,')]
```

```
In [ ]: # 4. Find the average Length of the movies of each genre.
df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
result_df = df.groupBy('Genre').agg(mean('Length').alias('AverageLength'))

# Show the results
result_df.show(truncate=False)
```

Genre	AverageLength
Crime	66.0
Romance	127.0
Adventure	119.0
NULL	120.5
Drama	113.3045525902668
War	116.90625
Fantasy	102.0
Mystery	103.00990099009901
Music	100.48780487804878
Science Fiction	106.47368421052632
CAT	NULL
Horror	93.92727272727272
Short	40.0
Western	93.0091743119266
Comedy	96.50540540540541
Action	104.5
Westerns	124.8


```
In [ ]: # 5. Find the Lead actor and Lead actress pair who have acted in more than one movie together.
df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
df.groupBy('Actor','Actress').agg(count('Title').alias('Count')).filter(col('Count')>1).collect()
```

```
Out[ ]: [Row(Actor=None, Actress='Madonna', Count=6),
Row(Actor='Schwarzenegger, Arnold', Actress='Hamilton, Linda', Count=2),
Row(Actor='Taylor, Robert', Actress=None, Count=3),
Row(Actor='Grant, Cary', Actress='Bergman, Ingrid', Count=2),
Row(Actor='Björnstrand, Gunnar', Actress='Andersson, Harriet', Count=3),
Row(Actor='De Niro, Robert', Actress='Streep, Meryl', Count=2),
Row(Actor='Pullman, Bill', Actress='Schaal, Wendy', Count=2),
Row(Actor='Cooper, Gary', Actress=None, Count=2),
Row(Actor='Van Damme, Jean-Claude', Actress=None, Count=2),
Row(Actor='Chapman, Graham', Actress=None, Count=6),
Row(Actor='Trintignant, Jean-Louis', Actress='Aimee, Anouk', Count=3),
Row(Actor='Cleese, John', Actress='Scales, Prunella', Count=2),
Row(Actor='Taylor, Rod', Actress=None, Count=3),
Row(Actor='Nicholson, Jack', Actress='Streep, Meryl', Count=2),
Row(Actor='Dean, Eddie', Actress=None, Count=2),
Row(Actor='Depardieu, Gérard', Actress=None, Count=4),
Row(Actor=None, Actress=None, Count=2),
Row(Actor='Shatner, William', Actress='Nichols, Nichelle', Count=2),
Row(Actor='Douglas, Michael', Actress='Turner, Kathleen', Count=2),
Row(Actor='Moore, Roger', Actress=None, Count=5),
Row(Actor='Ford, Harrison', Actress=None, Count=2),
Row(Actor='LaRue, Lash', Actress=None, Count=5),
Row(Actor='Eastwood, Clint', Actress='Locke, Sondra', Count=2),
Row(Actor='Allen, Woody', Actress='Farrow, Mia', Count=2),
Row(Actor='Wayne, John', Actress='Trevor, Claire', Count=2),
Row(Actor='Eastwood, Clint', Actress=None, Count=14),
Row(Actor='Douglas, Melvyn', Actress='Garbo, Greta', Count=2),
Row(Actor='Cleese, John', Actress=None, Count=5),
Row(Actor='Sydow, Max von', Actress='Ullman, Liv', Count=2),
Row(Actor='Wayne, John', Actress='O'Hara, Maureen', Count=4),
Row(Actor='Connery, Sean', Actress=None, Count=7),
Row(Actor='Gilbert, John', Actress='Garbo, Greta', Count=3),
Row(Actor='Slaughter, Tod', Actress=None, Count=2),
Row(Actor='Redford, Robert', Actress=None, Count=2),
Row(Actor='Montand, Yves', Actress='Deneuve, Catherine', Count=2),
Row(Actor='Brando, Marlon', Actress=None, Count=5),
Row(Actor='Quinn, Anthony', Actress='Loren, Sophia', Count=2),
Row(Actor='Scott, George C.', Actress=None, Count=2),
Row(Actor='Dillon, Matt', Actress='Lane, Diane', Count=2),
Row(Actor='Piscopo, Joe', Actress=None, Count=3),
Row(Actor='Allen, Woody', Actress='Keaton, Diane', Count=5),
Row(Actor='Harris, Richard', Actress=None, Count=2),
Row(Actor='Ford, Wallace', Actress=None, Count=2),
Row(Actor='Brando, Marlon', Actress='Simmons, Jean', Count=2),
Row(Actor='Nicholson, Jack', Actress=None, Count=3),
Row(Actor='Eastwood, Clint', Actress='Bloom, Verna', Count=2),
Row(Actor='Depardieu, Gérard', Actress='Deneuve, Catherine', Count=2),
Row(Actor='Karloff, Boris', Actress=None, Count=2),
Row(Actor='Gere, Richard', Actress='Basinger, Kim', Count=2),
Row(Actor='Murphy, Eddie', Actress=None, Count=5),
Row(Actor='Banderas, Antonio', Actress='Maura, Carmen', Count=2),
Row(Actor='Newman, Paul', Actress='Woodward, Joanne', Count=3),
Row(Actor='Sellers, Peter', Actress=None, Count=12),
Row(Actor='Welles, Orson', Actress=None, Count=4),
Row(Actor='Ford, Glenn', Actress=None, Count=5),
Row(Actor='Hamill, Mark', Actress='Fisher, Carrie', Count=3),
Row(Actor='Sellers, Peter', Actress='MacLaine, Shirley', Count=2),
Row(Actor='Newman, Paul', Actress=None, Count=7),
Row(Actor='Chase, Chevy', Actress='Hawn, Goldie', Count=2),
Row(Actor='Brown, Johnny Mack', Actress='O'Day, Nell', Count=3),
Row(Actor='Wilder, Gene', Actress=None, Count=2),
Row(Actor='Hale, Monte', Actress=None, Count=4),
Row(Actor='Tracy, Spencer', Actress=None, Count=8),
Row(Actor='Martin, Steve', Actress=None, Count=4),
Row(Actor='Sutherland, Donald', Actress=None, Count=2),
Row(Actor='Ritter, Tex', Actress=None, Count=2),
Row(Actor='Burton, Richard', Actress='Taylor, Elizabeth', Count=3),
Row(Actor='Hutton, Timothy', Actress=None, Count=2),
Row(Actor='Lane, Allan', Actress=None, Count=3),
Row(Actor='Hopkins, Anthony', Actress=None, Count=3),
Row(Actor='De Niro, Robert', Actress=None, Count=2),
Row(Actor='Hogan, Paul', Actress='Kozlowski, Linda', Count=2),
Row(Actor='Stone, Lewis', Actress='Garbo, Greta', Count=2),
Row(Actor='Williams, Robin', Actress='Reed, Pamela', Count=2),
Row(Actor='Hopper, Dennis', Actress=None, Count=2),
Row(Actor='Tracy, Spencer', Actress='Hepburn, Katharine', Count=11),
Row(Actor='Ronet, Maurice', Actress='Bardot, Brigitte', Count=2),
Row(Actor='Klein-Rogge, Rudolf', Actress=None, Count=2),
Row(Actor='Reynolds, Burt', Actress='Welch, Raquel', Count=2),
Row(Actor='Sellers, Peter', Actress='Ekland, Britt', Count=2),
Row(Actor='Mastroianni, Marcello', Actress='Loren, Sophia', Count=3),
Row(Actor='Adolphson, Edwin', Actress='Bergman, Ingrid', Count=2),
Row(Actor='Russell, Kurt', Actress='Hawn, Goldie', Count=2),
Row(Actor='Mifune, Toshiro', Actress=None, Count=7),
Row(Actor='Williams, Robin', Actress=None, Count=6),
Row(Actor='Moore, Dudley', Actress=None, Count=2),
Row(Actor='Lancaster, Burt', Actress=None, Count=12),
Row(Actor='Wayne, John', Actress=None, Count=42),
Row(Actor='Ford, Glenn', Actress='Hayworth, Rita', Count=3),
Row(Actor='Eastwood, Clint', Actress='Mason, Marsha', Count=2),
Row(Actor='Hopkins, Anthony', Actress='Bancroft, Anne', Count=2)]
```

```
In [ ]: # 6. Find the names of directors who directed movies of both the mystery and drama genres.
df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
# Filter movies of Mystery genre
A = df.filter(col('Genre') == 'Mystery').select('Director').distinct()

# Filter movies of Drama genre
B = df.filter(col('Genre') == 'Drama').select('Director').distinct()

# Join on the Director column
result_df = A.join(B, 'Director', 'inner')

# Show the results
result_df.show(truncate=False)
```

```
+-----+
|Director|
+-----+
|Avildsen, John G.|
|Stern, Steven Hilliard|
|Kurosawa, Akira|
|Lang, Fritz|
|Eastwood, Clint|
|Cimino, Michael|
|Kubrick, Stanley|
|Lumet, Sidney|
|Welles, Orson|
|Kagan, Jeremy Paul|
|Locke, Sondra|
|Lynch, David|
|Levinson, Barry|
|Scorsese, Martin|
|Taylor, Jud|
|Polanski, Roman|
|Hitchcock, Alfred|
|Rafelson, Bob|
|Mackenzie, John|
|Pakula, Alan J.|
+-----+
only showing top 20 rows
```

```
In [ ]: # 6 (B). find names of directors and all the genres of movies they have directed
df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
result_df = df.groupBy('Director').agg(collect_set('Genre').alias('Genres'))

# Show the results
result_df.show(truncate=False)
```

```
+-----+-----+
|Director|Genres|
+-----+-----+
|NULL|[War, Action, Science Fiction, Comedy, Drama, Horror, Music, Westerns, Mystery, Western]|
|Abrahams, Jim|[Comedy]|
|Adolphson, Edvin|[Comedy]|
|Alda, Alan|[Comedy]|
|Aldrich, Robert|[Action, Comedy, Drama, Mystery, Western]|
|Allegret, Marc|[Comedy]|
|Allen, Craig|[Drama]|
|Allen, Woody|[Comedy, Drama]|
|Almodóvar, Pedro|[Comedy, Drama]|
|Altman, Robert|[Comedy, Drama]|
|Amateau, Rod|[Drama]|
|Amurri, Franco|[Comedy]|
|Anderson, Michael|[Mystery]|
|Anger, Kenneth|[Short]|
|Annakin, Ken|[War]|
|Annaud, Jean-Jacques|[Comedy, Drama]|
|Ansbaugh, David|[Drama]|
|Anthony, Joseph|[Comedy, Drama]|
|Antonio, Lou|[Drama]|
|Antonioni, Michelangelo|[Drama]|
+-----+-----+
only showing top 20 rows
```

```
In [ ]: # 7. Find each director and find the average, max, and min ranking of his movies.  
df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')  
df.groupBy('Director').agg(mean('Popularity').alias('Average'),max('Popularity'),min('Popularity')).collect()
```

```

Out[ ]: [Row(Director=None, Average=40.569169960474305, max(Popularity)=9', min(Popularity)=0'),
Row(Director='Abrahams, Jim', Average=62.5, max(Popularity)=84', min(Popularity)=41'),
Row(Director='Adolphson, Edwin', Average=72.0, max(Popularity)=72', min(Popularity)=72'),
Row(Director='Alda, Alan', Average=32.5, max(Popularity)=53', min(Popularity)=12'),
Row(Director='Aldrich, Robert', Average=62.2, max(Popularity)=84', min(Popularity)=10'),
Row(Director='Allegret, Marc', Average=25.0, max(Popularity)=25', min(Popularity)=25'),
Row(Director='Allen, Craig', Average=None, max(Popularity)=None, min(Popularity)=None),
Row(Director='Allen, Woody', Average=37.625, max(Popularity)=82', min(Popularity)=14'),
Row(Director='Almodóvar, Pedro', Average=62.33333333333336, max(Popularity)=73', min(Popularity)=34'),
Row(Director='Altman, Robert', Average=11.0, max(Popularity)=3', min(Popularity)=19'),
Row(Director='Amateau, Rod', Average=80.0, max(Popularity)=80', min(Popularity)=80'),
Row(Director='Amurri, Franco', Average=19.0, max(Popularity)=19', min(Popularity)=19'),
Row(Director='Anderson, Michael', Average=34.0, max(Popularity)=34', min(Popularity)=34'),
Row(Director='Anger, Kenneth', Average=62.0, max(Popularity)=62', min(Popularity)=62'),
Row(Director='Annakin, Ken', Average=7.0, max(Popularity)=7', min(Popularity)=7'),
Row(Director='Annaud, Jean-Jacques', Average=16.0, max(Popularity)=8', min(Popularity)=24'),
Row(Director='Anspaugh, David', Average=2.0, max(Popularity)=2', min(Popularity)=2'),
Row(Director='Anthony, Joseph', Average=44.0, max(Popularity)=67', min(Popularity)=21'),
Row(Director='Antonio, Lou', Average=54.0, max(Popularity)=54', min(Popularity)=54'),
Row(Director='Antonioni, Michelangelo', Average=32.0, max(Popularity)=32', min(Popularity)=32'),
Row(Director='Apted, Michael', Average=32.0, max(Popularity)=41', min(Popularity)=23'),
Row(Director='Arkush, Allan', Average=34.0, max(Popularity)=34', min(Popularity)=34'),
Row(Director='Arnold, Newt', Average=78.0, max(Popularity)=78', min(Popularity)=78'),
Row(Director='Ashby, Hal', Average=50.0, max(Popularity)=69', min(Popularity)=31'),
Row(Director='Aslanian, Samson', Average=8.0, max(Popularity)=8', min(Popularity)=8'),
Row(Director='Attenborough, Richard', Average=54.33333333333336, max(Popularity)=85', min(Popularity)=7'),
Row(Director='Auer, John H.', Average=36.0, max(Popularity)=36', min(Popularity)=36'),
Row(Director='August, Bille', Average=14.0, max(Popularity)=14', min(Popularity)=14'),
Row(Director='Aurel, Jean', Average=66.0, max(Popularity)=66', min(Popularity)=66'),
Row(Director='Avedis, Howard Hikmet', Average=74.0, max(Popularity)=74', min(Popularity)=74'),
Row(Director='Avildsen, John G.', Average=70.33333333333333, max(Popularity)=82', min(Popularity)=51'),
Row(Director='Avnet, John', Average=52.0, max(Popularity)=52', min(Popularity)=52'),
Row(Director='Axel, Gabriel', Average=79.0, max(Popularity)=79', min(Popularity)=79'),
Row(Director='Aykroyd, Dan', Average=25.0, max(Popularity)=25', min(Popularity)=25'),
Row(Director='Babenco, Hector', Average=41.0, max(Popularity)=81', min(Popularity)=10'),
Row(Director='Badham, John', Average=37.5, max(Popularity)=62', min(Popularity)=13'),
Row(Director='Bahr, Fax', Average=72.0, max(Popularity)=72', min(Popularity)=72'),
Row(Director='Baker, Graham', Average=59.0, max(Popularity)=59', min(Popularity)=59'),
Row(Director='Bakshi, Ralph', Average=44.0, max(Popularity)=44', min(Popularity)=44'),
Row(Director='Band, Charles', Average=61.0, max(Popularity)=75', min(Popularity)=47'),
Row(Director='Barbash, Uri', Average=75.0, max(Popularity)=75', min(Popularity)=75'),
Row(Director='Barker, Clive', Average=72.0, max(Popularity)=72', min(Popularity)=72'),
Row(Director='Barnard, Michael', Average=5.0, max(Popularity)=5', min(Popularity)=5'),
Row(Director='Barreto, Bruno', Average=1.0, max(Popularity)=1', min(Popularity)=1'),
Row(Director='Barrett, Lezli-An', Average=17.0, max(Popularity)=17', min(Popularity)=17'),
Row(Director='Barry, Michael', Average=21.0, max(Popularity)=21', min(Popularity)=21'),
Row(Director='Battersby, Roy', Average=10.0, max(Popularity)=10', min(Popularity)=10'),
Row(Director='Beatty, Warren', Average=80.0, max(Popularity)=84', min(Popularity)=76'),
Row(Director='Beaudine, William', Average=83.0, max(Popularity)=83', min(Popularity)=83'),
Row(Director='Becker, Harold', Average=84.0, max(Popularity)=84', min(Popularity)=84'),
Row(Director='Beebe, Ford', Average=87.0, max(Popularity)=87', min(Popularity)=87'),
Row(Director='Beineix, Jean-Jacques', Average=50.0, max(Popularity)=71', min(Popularity)=29'),
Row(Director='Belson, Jerry', Average=84.0, max(Popularity)=84', min(Popularity)=84'),
Row(Director='Benedek, Laslo', Average=26.0, max(Popularity)=26', min(Popularity)=26'),
Row(Director='Benjamin, Richard', Average=50.0, max(Popularity)=50', min(Popularity)=50'),
Row(Director='Bennett, Walter', Average=47.0, max(Popularity)=47', min(Popularity)=47'),
Row(Director='Benton, Robert', Average=45.0, max(Popularity)=83', min(Popularity)=42'),
Row(Director='Beresford, Bruce', Average=57.75, max(Popularity)=84', min(Popularity)=6'),
Row(Director='Bergman, Andrew', Average=34.0, max(Popularity)=53', min(Popularity)=17'),
Row(Director='Bergman, Ingmar', Average=42.206896551724135, max(Popularity)=87', min(Popularity)=0'),
Row(Director='Bergman, Robert', Average=19.0, max(Popularity)=19', min(Popularity)=19'),
Row(Director='Bernds, Edward', Average=None, max(Popularity)=None, min(Popularity)=None),
Row(Director='Bernhardt, Curtis', Average=20.0, max(Popularity)=20', min(Popularity)=20'),
Row(Director='Berri, Claude', Average=63.5, max(Popularity)=87', min(Popularity)=40'),
Row(Director='Bertolucci, Bernardo', Average=36.0, max(Popularity)=72', min(Popularity)=1'),
Row(Director='Besson, Luc', Average=22.75, max(Popularity)=72', min(Popularity)=6'),
Row(Director='Betnel, Jonathan', Average=84.0, max(Popularity)=84', min(Popularity)=84'),
Row(Director='Bill, Tony', Average=67.0, max(Popularity)=73', min(Popularity)=61'),
Row(Director='Birch, Patricia', Average=64.0, max(Popularity)=64', min(Popularity)=64'),
Row(Director='Blancato, Ken', Average=28.0, max(Popularity)=28', min(Popularity)=28'),
Row(Director='Blier, Bertrand', Average=63.5, max(Popularity)=78', min(Popularity)=35'),
Row(Director='Blot, Philippe', Average=62.0, max(Popularity)=62', min(Popularity)=62'),
Row(Director='Bogart, Paul', Average=60.0, max(Popularity)=60', min(Popularity)=60'),
Row(Director='Bogdanovich, Peter', Average=32.5, max(Popularity)=62', min(Popularity)=3'),
Row(Director='Bogner, Willy', Average=72.0, max(Popularity)=72', min(Popularity)=72'),
Row(Director='Boisrond, Michel', Average=26.5, max(Popularity)=37', min(Popularity)=16'),
Row(Director='Boleslawski, Richard', Average=57.0, max(Popularity)=57', min(Popularity)=57'),
Row(Director='Bolt, Ben', Average=11.0, max(Popularity)=11', min(Popularity)=11'),
Row(Director='Bondarchuk, Sergei', Average=80.0, max(Popularity)=80', min(Popularity)=80'),
Row(Director='Bonerz, Peter', Average=29.0, max(Popularity)=29', min(Popularity)=29'),
Row(Director='Boorman, John', Average=12.666666666666666, max(Popularity)=6', min(Popularity)=29'),
Row(Director='Boulting, John', Average=38.0, max(Popularity)=38', min(Popularity)=38'),
Row(Director='Boulting, Roy', Average=52.0, max(Popularity)=63', min(Popularity)=41'),
Row(Director='Bourguignon, Serge', Average=32.5, max(Popularity)=54', min(Popularity)=11'),
Row(Director='Bowser, Kenneth', Average=72.0, max(Popularity)=72', min(Popularity)=72'),
Row(Director='Bradbury, Robert N.', Average=None, max(Popularity)=None, min(Popularity)=None),
Row(Director='Bradley, Al', Average=40.0, max(Popularity)=40', min(Popularity)=40'),
Row(Director='Brand, Larry', Average=9.0, max(Popularity)=9', min(Popularity)=9'),
Row(Director='Brando, Marlon', Average=26.0, max(Popularity)=26', min(Popularity)=26'),
Row(Director='Brest, Martin', Average=41.0, max(Popularity)=41', min(Popularity)=41'),
Row(Director='Brickman, Marshall', Average=51.0, max(Popularity)=51', min(Popularity)=51'),
Row(Director='Brickman, Paul', Average=47.0, max(Popularity)=66', min(Popularity)=28'),
Row(Director='Bridges, James', Average=26.666666666666668, max(Popularity)=7', min(Popularity)=30'),
Row(Director='Brooks, Albert', Average=75.0, max(Popularity)=75', min(Popularity)=75'),
Row(Director='Brooks, James L.', Average=32.0, max(Popularity)=32', min(Popularity)=32'),
Row(Director='Brooks, Mel', Average=30.0, max(Popularity)=33', min(Popularity)=27'),
Row(Director='Brooks, Richard', Average=60.0, max(Popularity)=66', min(Popularity)=54'),
Row(Director='Brown, Clarence', Average=51.2, max(Popularity)=83', min(Popularity)=0'),

```

```
Row(Director='Browning, Tod', Average=61.0, max(Popularity)=61, min(Popularity)=61),
Row(Director='Budds, Colin', Average=16.0, max(Popularity)=16, min(Popularity)=16),
Row(Director='Bunuel, Luis', Average=4.0, max(Popularity)=4, min(Popularity)=4),
Row(Director='Burton, Tim', Average=14.0, max(Popularity)=14, min(Popularity)=14),
Row(Director='Byrne, David', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Byrum, John', Average=69.5, max(Popularity)=85, min(Popularity)=54),
Row(Director='CAT', Average=None, max(Popularity)=INT, min(Popularity)=INT),
Row(Director='Cabanne, Christy', Average=33.0, max(Popularity)=33, min(Popularity)=33),
Row(Director='Cameron, James', Average=35.666666666666664, max(Popularity)=82, min(Popularity)=17),
Row(Director='Cammell, Donald', Average=48.0, max(Popularity)=48, min(Popularity)=48),
Row(Director='Cannon, Dyan', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Cannon, Raymond', Average=83.0, max(Popularity)=83, min(Popularity)=83),
Row(Director='Carlsen, Henning', Average=68.0, max(Popularity)=68, min(Popularity)=68),
Row(Director='Caro, Marc', Average=78.0, max(Popularity)=78, min(Popularity)=78),
Row(Director='Carpenter, John', Average=58.0, max(Popularity)=58, min(Popularity)=58),
Row(Director='Casden, Ron', Average=38.0, max(Popularity)=38, min(Popularity)=38),
Row(Director='Cassavetes, John', Average=60.0, max(Popularity)=60, min(Popularity)=60),
Row(Director='Castle, Nick', Average=25.0, max(Popularity)=25, min(Popularity)=25),
Row(Director='Caton-Jones, Michael', Average=64.0, max(Popularity)=64, min(Popularity)=64),
Row(Director='Chaffey, Don', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Chaplin, Charles', Average=4.0, max(Popularity)=4, min(Popularity)=4),
Row(Director='Chapman, Michael', Average=69.0, max(Popularity)=73, min(Popularity)=65),
Row(Director='Chelsom, Peter', Average=72.0, max(Popularity)=72, min(Popularity)=72),
Row(Director='Chomsky, Marvin J.', Average=1.0, max(Popularity)=1, min(Popularity)=1),
Row(Director='Chouraqui, Elie', Average=15.0, max(Popularity)=15, min(Popularity)=15),
Row(Director='Cimino, Michael', Average=49.25, max(Popularity)=82, min(Popularity)=16),
Row(Director='Clair, Rene', Average=60.0, max(Popularity)=60, min(Popularity)=60),
Row(Director='Clark, Bob', Average=36.0, max(Popularity)=36, min(Popularity)=36),
Row(Director='Clark, Jim', Average=31.0, max(Popularity)=31, min(Popularity)=31),
Row(Director='Clarke, Alan', Average=36.5, max(Popularity)=68, min(Popularity)=5),
Row(Director='Claxton, William F.', Average=28.0, max(Popularity)=28, min(Popularity)=28),
Row(Director='Clayton, Jack', Average=24.0, max(Popularity)=24, min(Popularity)=24),
Row(Director='Clegg, Tom', Average=45.0, max(Popularity)=45, min(Popularity)=45),
Row(Director='Clement, Dick', Average=47.0, max(Popularity)=47, min(Popularity)=47),
Row(Director='Clément, René', Average=63.0, max(Popularity)=63, min(Popularity)=63),
Row(Director='Cocteau, Jean', Average=14.0, max(Popularity)=14, min(Popularity)=14),
Row(Director='Coen, Joel', Average=23.0, max(Popularity)=23, min(Popularity)=23),
Row(Director='Cohen, Martin B.', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Colla, Richard A.', Average=22.0, max(Popularity)=7, min(Popularity)=37),
Row(Director='Collins, Lewis D.', Average=39.0, max(Popularity)=50, min(Popularity)=28),
Row(Director='Colmes, Walter', Average=26.0, max(Popularity)=26, min(Popularity)=26),
Row(Director='Comeau, Alain', Average=15.0, max(Popularity)=15, min(Popularity)=15),
Row(Director='Compton, Richard', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Conway, Jack', Average=86.0, max(Popularity)=86, min(Popularity)=86),
Row(Director='Cook, Fielder', Average=27.0, max(Popularity)=42, min(Popularity)=12),
Row(Director='Coolidge, Martha', Average=17.0, max(Popularity)=4, min(Popularity)=30),
Row(Director='Cooper, Jackie', Average=78.0, max(Popularity)=80, min(Popularity)=76),
Row(Director='Coppola, Christopher', Average=55.0, max(Popularity)=55, min(Popularity)=55),
Row(Director='Coppola, Francis Ford', Average=30.785714285714285, max(Popularity)=8, min(Popularity)=27),
Row(Director='Corman, Roger', Average=57.2, max(Popularity)=88, min(Popularity)=17),
Row(Director='Corneau, Alan', Average=87.0, max(Popularity)=87, min(Popularity)=87),
Row(Director='Cornell, John', Average=14.0, max(Popularity)=14, min(Popularity)=14),
Row(Director='Cosmatos, George P.', Average=57.0, max(Popularity)=57, min(Popularity)=57),
Row(Director='Costa-Gavras', Average=51.0, max(Popularity)=72, min(Popularity)=30),
Row(Director='Costner, Kevin', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Couturie, Bill', Average=57.0, max(Popularity)=57, min(Popularity)=57),
Row(Director='Cox, Alex', Average=47.0, max(Popularity)=47, min(Popularity)=47),
Row(Director='Crichton, Charles', Average=7.0, max(Popularity)=7, min(Popularity)=7),
Row(Director='Crichton, Michael', Average=44.333333333333336, max(Popularity)=7, min(Popularity)=62),
Row(Director='Cromwell, John', Average=2.0, max(Popularity)=2, min(Popularity)=2),
Row(Director='Cronenberg, David', Average=39.375, max(Popularity)=79, min(Popularity)=21),
Row(Director='Crosland, Alan', Average=55.0, max(Popularity)=55, min(Popularity)=55),
Row(Director='Cukor, George', Average=41.8, max(Popularity)=74, min(Popularity)=10),
Row(Director='Curtis, Dan', Average=48.5, max(Popularity)=62, min(Popularity)=35),
Row(Director='Curtiz, Michael', Average=10.0, max(Popularity)=10, min(Popularity)=10),
Row(Director='Dahlin, Bob', Average=39.0, max(Popularity)=39, min(Popularity)=39),
Row(Director='Dale, Cynthia', Average=63.0, max(Popularity)=63, min(Popularity)=63),
Row(Director='Damski, Mel', Average=34.0, max(Popularity)=34, min(Popularity)=34),
Row(Director='Daniel, Rod', Average=39.0, max(Popularity)=58, min(Popularity)=20),
Row(Director='Dante, Joe', Average=48.0, max(Popularity)=61, min(Popularity)=41),
Row(Director='Dassin, Jules', Average=39.0, max(Popularity)=39, min(Popularity)=39),
Row(Director='Daves, Delmer', Average=35.5, max(Popularity)=39, min(Popularity)=32),
Row(Director='Davis, Andrew', Average=24.0, max(Popularity)=24, min(Popularity)=24),
Row(Director='Day, Robert', Average=40.0, max(Popularity)=40, min(Popularity)=40),
Row(Director='De Bello, John', Average=35.5, max(Popularity)=47, min(Popularity)=24),
Row(Director='De Filippo, Eduardo', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='De Palma, Brian', Average=35.0, max(Popularity)=82, min(Popularity)=16),
Row(Director='De Sica, Vittorio', Average=55.666666666666664, max(Popularity)=83, min(Popularity)=36),
Row(Director='De Winter, Harry', Average=74.0, max(Popularity)=74, min(Popularity)=74),
Row(Director='DeCoteau, David', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='DeMille, Cecil B.', Average=59.0, max(Popularity)=59, min(Popularity)=59),
Row(Director='DeVito, Danny', Average=21.0, max(Popularity)=21, min(Popularity)=21),
Row(Director='Dearden, Basil', Average=49.5, max(Popularity)=75, min(Popularity)=24),
Row(Director='Degas, Andre', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Dell, Stewart', Average=9.0, max(Popularity)=9, min(Popularity)=9),
Row(Director='Dembo, Richard', Average=7.0, max(Popularity)=7, min(Popularity)=7),
Row(Director='Demme, Jonathan', Average=32.333333333333336, max(Popularity)=81, min(Popularity)=8),
Row(Director='Depardieu, Gérard', Average=67.0, max(Popularity)=67, min(Popularity)=67),
Row(Director='Deruddere, Dominique', Average=29.0, max(Popularity)=29, min(Popularity)=29),
Row(Director='Deutch, Howard', Average=41.333333333333336, max(Popularity)=75, min(Popularity)=16),
Row(Director='Di Leo, Mario', Average=10.0, max(Popularity)=10, min(Popularity)=10),
Row(Director='Dickerson, Ernest R.', Average=31.0, max(Popularity)=31, min(Popularity)=31),
Row(Director='Dinner, Michael', Average=78.0, max(Popularity)=78, min(Popularity)=78),
Row(Director='Dmytryk, Edward', Average=28.0, max(Popularity)=74, min(Popularity)=0),
Row(Director='Donaggio, Michael Winner', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Donaldson, Roger', Average=26.0, max(Popularity)=38, min(Popularity)=13),
Row(Director='Donen, Stanley', Average=27.25, max(Popularity)=67, min(Popularity)=1),
Row(Director='Donner, Clive', Average=64.5, max(Popularity)=83, min(Popularity)=46),
Row(Director='Donner, Richard', Average=59.75, max(Popularity)=87, min(Popularity)=15),
```

```
Row(Director='Douglas, Kirk', Average=76.0, max(Popularity)=76, min(Popularity)=76),
Row(Director='Dragoti, Stan', Average=60.0, max(Popularity)=60, min(Popularity)=60),
Row(Director='Drury, David', Average=52.0, max(Popularity)=52, min(Popularity)=52),
Row(Director='Duigan, John', Average=71.0, max(Popularity)=71, min(Popularity)=71),
Row(Director='Dwan, Allan', Average=62.0, max(Popularity)=72, min(Popularity)=52),
Row(Director='Eastwood, Clint', Average=51.0, max(Popularity)=77, min(Popularity)=12),
Row(Director='Edgren, Gustaf', Average=32.0, max(Popularity)=32, min(Popularity)=32),
Row(Director='Edwards, Blake', Average=54.57142857142857, max(Popularity)=88, min(Popularity)=32),
Row(Director='Edzard, Christine', Average=12.0, max(Popularity)=12, min(Popularity)=12),
Row(Director='Emes, Ian', Average=None, max(Popularity)=None, min(Popularity)=None),
Row(Director='Englund, George', Average=63.0, max(Popularity)=63, min(Popularity)=63),
Row(Director='Estevez, Emilio', Average=25.0, max(Popularity)=25, min(Popularity)=25),
Row(Director='Faiman, Peter', Average=66.0, max(Popularity)=66, min(Popularity)=66),
Row(Director='Fanaka, Jamaa', Average=82.0, max(Popularity)=82, min(Popularity)=82),
Row(Director='Farrow, John', Average=4.0, max(Popularity)=4, min(Popularity)=4),
Row(Director='Feferman, Linda', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Fellini, Federico', Average=58.857142857142854, max(Popularity)=88, min(Popularity)=20),
Row(Director='Ferrera, Abel', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Figgis, Mike', Average=3.0, max(Popularity)=3, min(Popularity)=3),
Row(Director='Finch, Charles', Average=7.0, max(Popularity)=7, min(Popularity)=7),
Row(Director='Fincher, David', Average=59.0, max(Popularity)=59, min(Popularity)=59),
Row(Director='Finkleman, Ken', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Fisher, Jack', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Fitzmaurice, George', Average=67.0, max(Popularity)=67, min(Popularity)=67),
Row(Director='Fleischer, Richard', Average=86.0, max(Popularity)=86, min(Popularity)=86),
Row(Director='Fleming, Victor', Average=24.5, max(Popularity)=7, min(Popularity)=42),
Row(Director='Flicker, Theodore J.', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Florea, John', Average=0.0, max(Popularity)=0, min(Popularity)=0),
Row(Director='Foley, James', Average=23.5, max(Popularity)=33, min(Popularity)=14),
Row(Director='Forbes, Bryan', Average=40.0, max(Popularity)=40, min(Popularity)=40),
Row(Director='Ford, Aleksander', Average=73.0, max(Popularity)=73, min(Popularity)=73),
Row(Director='Ford, Greg', Average=68.0, max(Popularity)=68, min(Popularity)=68),
Row(Director='Ford, John', Average=47.958333333333336, max(Popularity)=9, min(Popularity)=1),
Row(Director='Ford, Philip', Average=28.142857142857142, max(Popularity)=8, min(Popularity)=1),
Row(Director='Forman, Milos', Average=27.0, max(Popularity)=8, min(Popularity)=6),
Row(Director='Forsyth, Bill', Average=54.0, max(Popularity)=54, min(Popularity)=54),
Row(Director='Fosse, Bob', Average=59.0, max(Popularity)=59, min(Popularity)=59),
Row(Director='Fraker, William A.', Average=29.0, max(Popularity)=29, min(Popularity)=29),
Row(Director='Frank, Melvin', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Frankenheimer, John', Average=41.333333333333336, max(Popularity)=9, min(Popularity)=39),
Row(Director='Franklin, Howard', Average=24.0, max(Popularity)=24, min(Popularity)=24),
Row(Director='Franklin, Sidney', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='Frears, Stephen', Average=41.5, max(Popularity)=77, min(Popularity)=6),
Row(Director='Freed, Mark', Average=31.0, max(Popularity)=31, min(Popularity)=31),
Row(Director='Friedkin, William', Average=62.0, max(Popularity)=88, min(Popularity)=28),
Row(Director='Fukasaku, Kinji', Average=62.0, max(Popularity)=62, min(Popularity)=62),
Row(Director='Furie, Sidney J.', Average=46.0, max(Popularity)=77, min(Popularity)=15),
Row(Director='Garnett, Tay', Average=24.0, max(Popularity)=24, min(Popularity)=24),
Row(Director='George, Peter', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Gibbins, Duncan', Average=9.0, max(Popularity)=9, min(Popularity)=9),
Row(Director='Gibson, Alan', Average=15.0, max(Popularity)=15, min(Popularity)=15),
Row(Director='Gilbert, Brian', Average=42.0, max(Popularity)=55, min(Popularity)=29),
Row(Director='Gilbert, Lewis', Average=46.75, max(Popularity)=77, min(Popularity)=27),
Row(Director='Gillard, Stuart', Average=72.0, max(Popularity)=72, min(Popularity)=72),
Row(Director='Gilliam, Terry', Average=32.0, max(Popularity)=83, min(Popularity)=5),
Row(Director='Glaser, Paul Michael', Average=31.0, max(Popularity)=31, min(Popularity)=31),
Row(Director='Glatter, Leslie Linka', Average=34.0, max(Popularity)=34, min(Popularity)=34),
Row(Director='Glen, John', Average=64.33333333333333, max(Popularity)=86, min(Popularity)=39),
Row(Director='Godard, Jean-Luc', Average=50.5, max(Popularity)=81, min(Popularity)=20),
Row(Director='Gold, Greg', Average=45.0, max(Popularity)=45, min(Popularity)=45),
Row(Director='Golden, Dan', Average=26.0, max(Popularity)=26, min(Popularity)=26),
Row(Director='Gordon, Steve', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Gottlieb, Michael', Average=23.0, max(Popularity)=23, min(Popularity)=23),
Row(Director='Goulding, Edmund', Average=49.5, max(Popularity)=81, min(Popularity)=18),
Row(Director='Gower, Bruce', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Graef, Roger', Average=36.0, max(Popularity)=36, min(Popularity)=36),
Row(Director='Grant, James Edward', Average=84.0, max(Popularity)=84, min(Popularity)=84),
Row(Director='Green, Alfred E.', Average=66.0, max(Popularity)=66, min(Popularity)=66),
Row(Director='Green, Guy', Average=26.5, max(Popularity)=51, min(Popularity)=2),
Row(Director='Greenaway, Peter', Average=15.5, max(Popularity)=3, min(Popularity)=28),
Row(Director='Gries, Tom', Average=38.0, max(Popularity)=48, min(Popularity)=28),
Row(Director='Griffi, Giuseppe Patroni', Average=72.0, max(Popularity)=72, min(Popularity)=72),
Row(Director='Grosbard, Ulu', Average=31.0, max(Popularity)=31, min(Popularity)=31),
Row(Director='Guest, Val', Average=54.0, max(Popularity)=54, min(Popularity)=54),
Row(Director='Guillermo, John', Average=20.0, max(Popularity)=20, min(Popularity)=20),
Row(Director='Guney, Yilmaz', Average=53.0, max(Popularity)=53, min(Popularity)=53),
Row(Director='Hackford, Taylor', Average=31.5, max(Popularity)=62, min(Popularity)=1),
Row(Director='Haggard, Piers', Average=29.0, max(Popularity)=29, min(Popularity)=29),
Row(Director='Haines, Fred', Average=20.0, max(Popularity)=20, min(Popularity)=20),
Row(Director='Haines, Randa', Average=20.0, max(Popularity)=20, min(Popularity)=20),
Row(Director='Hallström, Lasse', Average=44.5, max(Popularity)=68, min(Popularity)=21),
Row(Director='Hamilton, Guy', Average=42.0, max(Popularity)=8, min(Popularity)=41),
Row(Director='Hancock, John D.', Average=73.0, max(Popularity)=73, min(Popularity)=73),
Row(Director='Hanson, Curtis', Average=4.0, max(Popularity)=4, min(Popularity)=4),
Row(Director='Hardy, Rod', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Harrington, Curtis', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Harvey, Anthony', Average=78.0, max(Popularity)=78, min(Popularity)=78),
Row(Director='Haskin, Byron', Average=29.5, max(Popularity)=56, min(Popularity)=3),
Row(Director='Hathaway, Henry', Average=55.25, max(Popularity)=84, min(Popularity)=29),
Row(Director='Heckerling, Amy', Average=34.0, max(Popularity)=65, min(Popularity)=3),
Row(Director='Heisler, Stuart', Average=23.0, max(Popularity)=23, min(Popularity)=23),
Row(Director='Hellman, Monte', Average=48.0, max(Popularity)=70, min(Popularity)=26),
Row(Director='Hellman, Oliver', Average=62.0, max(Popularity)=62, min(Popularity)=62),
Row(Director='Herman, Mark', Average=69.0, max(Popularity)=69, min(Popularity)=69),
Row(Director='Hermoso, Miguel', Average=45.0, max(Popularity)=45, min(Popularity)=45),
Row(Director='Heston, Charlton', Average=40.0, max(Popularity)=40, min(Popularity)=40),
Row(Director='Heyes, Douglas', Average=26.0, max(Popularity)=26, min(Popularity)=26),
Row(Director='Hickox, Douglas', Average=64.0, max(Popularity)=64, min(Popularity)=64),
Row(Director='Higgin, Howard', Average=2.0, max(Popularity)=2, min(Popularity)=2),
```

```
Row(Director='Higgins, Colin', Average=46.0, max(Popularity)='46', min(Popularity)='46'),
Row(Director='Hill, George Roy', Average=48.5, max(Popularity)='83', min(Popularity)='29'),
Row(Director='Hill, Walter', Average=66.0, max(Popularity)='67', min(Popularity)='65'),
Row(Director='Hitchcock, Alfred', Average=42.44117647058823, max(Popularity)='8', min(Popularity)='10'),
Row(Director='Hobbs, Lyndall', Average=45.0, max(Popularity)='45', min(Popularity)='45'),
Row(Director='Hoffs, Tamar Simon', Average=71.0, max(Popularity)='71', min(Popularity)='71'),
Row(Director='Holcomb, Rod', Average=27.0, max(Popularity)='27', min(Popularity)='27'),
Row(Director='Holland, Savage Steve', Average=61.0, max(Popularity)='61', min(Popularity)='61'),
Row(Director='Holzman, Allan', Average=47.0, max(Popularity)='71', min(Popularity)='23'),
Row(Director='Hook, Harry', Average=41.0, max(Popularity)='41', min(Popularity)='41'),
Row(Director='Hooks, Kevin', Average=3.0, max(Popularity)='3', min(Popularity)='3'),
Row(Director='Hopkins, Stephen', Average=60.0, max(Popularity)='79', min(Popularity)='41'),
Row(Director='Hopper, Dennis', Average=31.0, max(Popularity)='70', min(Popularity)='22'),
Row(Director='Houck, Joy, Jr.', Average=88.0, max(Popularity)='88', min(Popularity)='88'),
Row(Director='Hough, John', Average=53.0, max(Popularity)='53', min(Popularity)='53'),
Row(Director='Howard, Ron', Average=52.0, max(Popularity)='59', min(Popularity)='45'),
Row(Director='Howard, William K.', Average=83.0, max(Popularity)='83', min(Popularity)='83'),
Row(Director='Howson, Frank', Average=68.0, max(Popularity)='68', min(Popularity)='68'),
Row(Director='Hudson, Hugh', Average=6.0, max(Popularity)='6', min(Popularity)='6'),
Row(Director='Hughes, John', Average=29.0, max(Popularity)='73', min(Popularity)='12'),
Row(Director='Hughes, Ken', Average=11.0, max(Popularity)='11', min(Popularity)='11'),
Row(Director='Hughes, Terry', Average=81.0, max(Popularity)='81', min(Popularity)='81'),
Row(Director='Hulton, Brian G.', Average=57.0, max(Popularity)='57', min(Popularity)='57'),
Row(Director='Hunt, Peter R.', Average=33.0, max(Popularity)='66', min(Popularity)='0'),
Row(Director='Hunter, Tim', Average=3.0, max(Popularity)='3', min(Popularity)='3'),
Row(Director='Huston, John', Average=49.111111111111114, max(Popularity)='81', min(Popularity)='25'),
Row(Director='Hutton, Brian G.', Average=85.5, max(Popularity)='87', min(Popularity)='84'),
Row(Director='Huyck, Willard', Average=75.0, max(Popularity)='75', min(Popularity)='75'),
Row(Director='Hyam, Peter', Average=3.0, max(Popularity)='3', min(Popularity)='3'),
Row(Director='Hyams, Peter', Average=30.666666666666668, max(Popularity)='81', min(Popularity)='4'),
Row(Director='Imamura, Shohei', Average=88.0, max(Popularity)='88', min(Popularity)='88'),
Row(Director='Ingvordsen, J. Christian', Average=66.0, max(Popularity)='66', min(Popularity)='66'),
Row(Director='Irvin, John', Average=63.0, max(Popularity)='63', min(Popularity)='63'),
Row(Director='Ivory, James', Average=60.0, max(Popularity)='75', min(Popularity)='45'),
Row(Director='Jackson, Donald G.', Average=52.5, max(Popularity)='74', min(Popularity)='31'),
Row(Director='Jackson, Mick', Average=30.0, max(Popularity)='30', min(Popularity)='30'),
Row(Director='Jaglom, Henry', Average=88.0, max(Popularity)='88', min(Popularity)='88'),
Row(Director='Jameson, Jerry', Average=39.5, max(Popularity)='69', min(Popularity)='10'),
Row(Director='Jarman, Derek', Average=41.0, max(Popularity)='41', min(Popularity)='41'),
Row(Director='Jarmusch, Jim', Average=32.0, max(Popularity)='49', min(Popularity)='23'),
Row(Director='Jeunet, Jean-Pierre', Average=60.0, max(Popularity)='60', min(Popularity)='60'),
Row(Director='Jewison, Norman', Average=58.75, max(Popularity)='79', min(Popularity)='6'),
Row(Director='Joanou, Phil', Average=49.5, max(Popularity)='50', min(Popularity)='49'),
Row(Director='Joffe, Roland', Average=49.75, max(Popularity)='87', min(Popularity)='20'),
Row(Director='Jones, Chris', Average=25.0, max(Popularity)='25', min(Popularity)='25'),
Row(Director='Jones, David', Average=9.0, max(Popularity)='9', min(Popularity)='9'),
Row(Director='Jones, Terry', Average=23.0, max(Popularity)='33', min(Popularity)='11'),
Row(Director='Jordan, Glenn', Average=48.0, max(Popularity)='48', min(Popularity)='48'),
Row(Director='Jordan, Neil', Average=52.0, max(Popularity)='53', min(Popularity)='51'),
Row(Director='Kaczender, George', Average=71.0, max(Popularity)='71', min(Popularity)='71'),
Row(Director='Kadar, Jan', Average=37.0, max(Popularity)='37', min(Popularity)='37'),
Row(Director='Kagan, Jeremy Paul', Average=18.0, max(Popularity)='19', min(Popularity)='17'),
Row(Director='Kalatozov, Mikhail', Average=7.0, max(Popularity)='7', min(Popularity)='7'),
Row(Director='Kampmann, Steven', Average=76.0, max(Popularity)='76', min(Popularity)='76'),
Row(Director='Kane, Joseph', Average=40.5, max(Popularity)='54', min(Popularity)='27'),
Row(Director='Kaplan, Jonathan', Average=71.0, max(Popularity)='71', min(Popularity)='71'),
Row(Director='Karlson, Phil', Average=60.0, max(Popularity)='60', min(Popularity)='60'),
Row(Director='Katselas, Milton', Average=82.0, max(Popularity)='82', min(Popularity)='82'),
Row(Director='Kaufman, Philip', Average=5.0, max(Popularity)='5', min(Popularity)='5'),
Row(Director='Kaurismäki, Mika', Average=30.0, max(Popularity)='30', min(Popularity)='30'),
Row(Director='Kazan, Elia', Average=78.66666666666667, max(Popularity)='86', min(Popularity)='75'),
Row(Director='Keach, James', Average=69.0, max(Popularity)='69', min(Popularity)='69'),
Row(Director='Kelllogg, Ray', Average=36.0, max(Popularity)='36', min(Popularity)='36'),
Row(Director='Kennedy, Burt', Average=84.0, max(Popularity)='84', min(Popularity)='84'),
Row(Director='Kershner, Irvin', Average=15.666666666666666, max(Popularity)='8', min(Popularity)='33'),
Row(Director='Keshishian, Alek', Average=54.0, max(Popularity)='54', min(Popularity)='54'),
Row(Director='Kiersch, Fritz', Average=2.0, max(Popularity)='2', min(Popularity)='2'),
Row(Director='Kikoine, Gerard', Average=71.0, max(Popularity)='71', min(Popularity)='71'),
Row(Director='King Keller, Frederick', Average=70.0, max(Popularity)='70', min(Popularity)='70'),
Row(Director='King, George', Average=45.0, max(Popularity)='45', min(Popularity)='45'),
Row(Director='King, Henry', Average=11.0, max(Popularity)='11', min(Popularity)='11'),
Row(Director='King, Rick', Average=41.0, max(Popularity)='41', min(Popularity)='41'),
Row(Director='King, Stephen', Average=40.0, max(Popularity)='40', min(Popularity)='40'),
Row(Director='Kleiser, Randal', Average=17.0, max(Popularity)='17', min(Popularity)='17'),
Row(Director='Kloves, Steve', Average=66.0, max(Popularity)='66', min(Popularity)='66'),
Row(Director='Kollek, Amos', Average=69.0, max(Popularity)='69', min(Popularity)='69'),
Row(Director='Konchalovsky, Andrei', Average=32.666666666666664, max(Popularity)='9', min(Popularity)='7'),
Row(Director='Koster, Henry', Average=54.75, max(Popularity)='84', min(Popularity)='22'),
Row(Director='Kotchhoff, Ted', Average=47.0, max(Popularity)='64', min(Popularity)='30'),
Row(Director='Kramer, Stanley', Average=44.5, max(Popularity)='50', min(Popularity)='39'),
Row(Director='Kronsberg, Jeremy Joe', Average=65.0, max(Popularity)='65', min(Popularity)='65'),
Row(Director='Kubrick, Stanley', Average=61.55555555555556, max(Popularity)='83', min(Popularity)='32'),
Row(Director='Kulik, Buzz', Average=51.0, max(Popularity)='51', min(Popularity)='51'),
Row(Director='Kurosawa, Akira', Average=52.14285714285715, max(Popularity)='9', min(Popularity)='31'),
Row(Director='Kuzui, Fran Rubel', Average=2.0, max(Popularity)='2', min(Popularity)='2'),
Row(Director='Lafia, John', Average=65.0, max(Popularity)='65', min(Popularity)='65'),
Row(Director='Lambert, Bruce', Average=17.25, max(Popularity)='26', min(Popularity)='13'),
Row(Director='Lambert, Mary', Average=37.5, max(Popularity)='48', min(Popularity)='27'),
Row(Director='Landis, John', Average=36.5, max(Popularity)='62', min(Popularity)='11'),
Row(Director='Lang, Fritz', Average=47.72727272727273, max(Popularity)='80', min(Popularity)='27'),
Row(Director='Lang, Walter', Average=51.0, max(Popularity)='51', min(Popularity)='51'),
Row(Director='Lanza, Anthony M.', Average=27.0, max(Popularity)='27', min(Popularity)='27'),
Row(Director='Lathan, Stan', Average=72.0, max(Popularity)='72', min(Popularity)='72'),
Row(Director='Lattuada, Alberto', Average=69.0, max(Popularity)='69', min(Popularity)='69'),
Row(Director='Lawton, J.F.', Average=56.0, max(Popularity)='56', min(Popularity)='56'),
Row(Director='LeRoy, Mervyn', Average=57.0, max(Popularity)='69', min(Popularity)='45'),
Row(Director='Lean, David', Average=81.0, max(Popularity)='81', min(Popularity)='81'),
Row(Director='Lederman, Ross', Average=51.0, max(Popularity)='51', min(Popularity)='51'),
Row(Director='Lee, Spike', Average=33.666666666666664, max(Popularity)='78', min(Popularity)='18'),
```

```
Row(Director='Leigh, Jason', Average=12.0, max(Popularity)=12, min(Popularity)=12),
Row(Director='Lelouch, Claude', Average=33.666666666666664, max(Popularity)=6, min(Popularity)=46),
Row(Director='Lennon, Biddy W.', Average=35.0, max(Popularity)=35, min(Popularity)=35),
Row(Director='Leo, Malcolm', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Leonard, Robert Z.', Average=64.0, max(Popularity)=64, min(Popularity)=64),
Row(Director='Leone, Sergio', Average=47.5, max(Popularity)=61, min(Popularity)=34),
Row(Director='Lester, Mark L.', Average=23.0, max(Popularity)=23, min(Popularity)=23),
Row(Director='Lester, Richard', Average=25.0, max(Popularity)=63, min(Popularity)=6),
Row(Director='Levinson, Barry', Average=27.75, max(Popularity)=8, min(Popularity)=16),
Row(Director='Levy, Ralph', Average=7.0, max(Popularity)=7, min(Popularity)=7),
Row(Director='Lewicki, Steven Jon', Average=24.0, max(Popularity)=24, min(Popularity)=24),
Row(Director='Lewis, Joseph H.', Average=35.0, max(Popularity)=35, min(Popularity)=35),
Row(Director='Lindberg, Per', Average=14.0, max(Popularity)=14, min(Popularity)=14),
Row(Director='Lisberger, Steven', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Litvak, Anatole', Average=15.0, max(Popularity)=6, min(Popularity)=24),
Row(Director='Lloyd, Frank', Average=76.0, max(Popularity)=76, min(Popularity)=76),
Row(Director='Loach, Ken', Average=71.0, max(Popularity)=71, min(Popularity)=71),
Row(Director='Locke, Sondra', Average=12.0, max(Popularity)=23, min(Popularity)=1),
Row(Director='Logan, Joshua', Average=27.0, max(Popularity)=46, min(Popularity)=16),
Row(Director='Loma, José Antonio de la', Average=14.0, max(Popularity)=14, min(Popularity)=14),
Row(Director='Lommel, Ulli', Average=17.0, max(Popularity)=17, min(Popularity)=17),
Row(Director='Loncraine, Richard', Average=15.0, max(Popularity)=15, min(Popularity)=15),
Row(Director='London, Roy', Average=67.0, max(Popularity)=67, min(Popularity)=67),
Row(Director='Losey, Joseph', Average=60.0, max(Popularity)=60, min(Popularity)=60),
Row(Director='Lovy, Steven', Average=78.0, max(Popularity)=78, min(Popularity)=78),
Row(Director='Lubitsch, Ernst', Average=40.0, max(Popularity)=40, min(Popularity)=40),
Row(Director='Lucas, George', Average=41.5, max(Popularity)=44, min(Popularity)=39),
Row(Director='Lucidi, Maurizio', Average=25.0, max(Popularity)=25, min(Popularity)=25),
Row(Director='Ludman, Larry', Average=35.5, max(Popularity)=37, min(Popularity)=34),
Row(Director='Ludwig, Edward', Average=24.5, max(Popularity)=35, min(Popularity)=14),
Row(Director='Lumet, Sidney', Average=16.142857142857142, max(Popularity)=8, min(Popularity)=1),
Row(Director='Lynch, David', Average=22.5, max(Popularity)=86, min(Popularity)=2),
Row(Director='Lyne, Adrian', Average=34.0, max(Popularity)=7, min(Popularity)=61),
Row(Director='MacNaughton, Ian', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Mackendrick, Alexander', Average=20.0, max(Popularity)=28, min(Popularity)=12),
Row(Director='MacKenzie, John', Average=32.5, max(Popularity)=51, min(Popularity)=14),
Row(Director='Madsen, Kenneth', Average=76.0, max(Popularity)=76, min(Popularity)=76),
Row(Director='Magnoli, Albert', Average=74.0, max(Popularity)=74, min(Popularity)=74),
Row(Director='Malick, Terrence', Average=14.0, max(Popularity)=14, min(Popularity)=14),
Row(Director='Malle, Louis', Average=27.333333333333332, max(Popularity)=35, min(Popularity)=17),
Row(Director='Malmut, Bruce', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Mamoulian, Rouben', Average=82.0, max(Popularity)=82, min(Popularity)=82),
Row(Director='Mandel, Robert', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Mankiewicz, Joseph L.', Average=55.75, max(Popularity)=80, min(Popularity)=23),
Row(Director='Mann, Anthony', Average=36.0, max(Popularity)=62, min(Popularity)=10),
Row(Director='Mann, Daniel', Average=30.5, max(Popularity)=50, min(Popularity)=11),
Row(Director='Mann, Delbert', Average=13.0, max(Popularity)=13, min(Popularity)=13),
Row(Director='Mann, Michael', Average=19.0, max(Popularity)=19, min(Popularity)=19),
Row(Director='Manning, Michelle', Average=38.0, max(Popularity)=38, min(Popularity)=38),
Row(Director='Marder, Bruce', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Marquand, Richard', Average=4.0, max(Popularity)=4, min(Popularity)=4),
Row(Director='Marshall, Garry', Average=24.5, max(Popularity)=6, min(Popularity)=43),
Row(Director='Marshall, George', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Marshall, Penny', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Martino, Sergio', Average=33.0, max(Popularity)=33, min(Popularity)=33),
Row(Director='Masterson, Peter', Average=62.0, max(Popularity)=62, min(Popularity)=62),
Row(Director='Mastorakis, Nico', Average=0.0, max(Popularity)=0, min(Popularity)=0),
Row(Director='Mate, Rudolph', Average=39.5, max(Popularity)=54, min(Popularity)=25),
Row(Director='Matthau, Charles', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Mattoli, Mario', Average=54.0, max(Popularity)=54, min(Popularity)=54),
Row(Director='May, Elaine', Average=83.0, max(Popularity)=83, min(Popularity)=83),
Row(Director='Mazursky, Paul', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='McBride, Jim', Average=51.0, max(Popularity)=51, min(Popularity)=51),
Row(Director='McCarey, Leo', Average=31.0, max(Popularity)=31, min(Popularity)=31),
Row(Director='McEveety, Bernard', Average=33.0, max(Popularity)=33, min(Popularity)=33),
Row(Director='McEveety, Vincent', Average=69.0, max(Popularity)=69, min(Popularity)=69),
Row(Director='McGann, William', Average=27.0, max(Popularity)=27, min(Popularity)=27),
Row(Director='McGrath, Joseph', Average=57.0, max(Popularity)=86, min(Popularity)=28),
Row(Director='McLaglen, Andrew V.', Average=35.857142857142854, max(Popularity)=9, min(Popularity)=12),
Row(Director='McTiernan, John', Average=7.0, max(Popularity)=8, min(Popularity)=6),
Row(Director='Meagher, John', Average=82.0, max(Popularity)=82, min(Popularity)=82),
Row(Director='Menendez, Ramon', Average=19.0, max(Popularity)=19, min(Popularity)=19),
Row(Director='Menzel, Jiri', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Meyer, Nicholas', Average=12.5, max(Popularity)=14, min(Popularity)=11),
Row(Director='Miles, Christopher', Average=85.0, max(Popularity)=85, min(Popularity)=85),
Row(Director='Milius, John', Average=45.25, max(Popularity)=83, min(Popularity)=2),
Row(Director='Miller, David', Average=33.5, max(Popularity)=61, min(Popularity)=6),
Row(Director='Miller, George', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Miller, Robert Ellis', Average=12.0, max(Popularity)=23, min(Popularity)=11),
Row(Director='Miller, Stuart', Average=76.0, max(Popularity)=76, min(Popularity)=76),
Row(Director='Minnelli, Vincente', Average=62.8, max(Popularity)=71, min(Popularity)=52),
Row(Director='Mischer, Don', Average=68.0, max(Popularity)=68, min(Popularity)=68),
Row(Director='Mizoguchi, Kenji', Average=82.0, max(Popularity)=82, min(Popularity)=82),
Row(Director='Mizrahi, Moshe', Average=11.0, max(Popularity)=11, min(Popularity)=11),
Row(Director='Molander, Gustaf', Average=56.333333333333336, max(Popularity)=88, min(Popularity)=32),
Row(Director='Molander, Gustav', Average=22.5, max(Popularity)=26, min(Popularity)=19),
Row(Director='Molinaro, Edoard', Average=34.0, max(Popularity)=34, min(Popularity)=34),
Row(Director='Morahan, Christopher', Average=10.0, max(Popularity)=10, min(Popularity)=10),
Row(Director='Morneau, Louis', Average=53.0, max(Popularity)=53, min(Popularity)=53),
Row(Director='Morrisey, Paul', Average=68.0, max(Popularity)=68, min(Popularity)=68),
Row(Director='Moxey, John Llewellyn', Average=71.0, max(Popularity)=71, min(Popularity)=71),
Row(Director='Mulcahy, Russell', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Mulligan, Robert', Average=35.5, max(Popularity)=60, min(Popularity)=11),
Row(Director='Murphy, Eddie', Average=31.0, max(Popularity)=51, min(Popularity)=11),
Row(Director='Myerson, Alan', Average=59.0, max(Popularity)=59, min(Popularity)=59),
Row(Director='Neame, Ronald', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Needham, Hal', Average=77.5, max(Popularity)=80, min(Popularity)=75),
Row(Director='Neff, Thomas L.', Average=63.0, max(Popularity)=63, min(Popularity)=63),
Row(Director='Negulesco, Jean', Average=65.0, max(Popularity)=65, min(Popularity)=65),
```



```

Row(Director='Nelson, Gary', Average=47.0, max(Popularity)=47, min(Popularity)=47),
Row(Director='Nelson, Ralph', Average=38.0, max(Popularity)=38, min(Popularity)=38),
Row(Director='Newell, Mike', Average=42.0, max(Popularity)=42, min(Popularity)=42),
Row(Director='Newman, Paul', Average=41.0, max(Popularity)=7, min(Popularity)=32),
Row(Director='Niblo, Fred', Average=72.0, max(Popularity)=72, min(Popularity)=72),
Row(Director='Nichols, Mike', Average=48.42857142857143, max(Popularity)=82, min(Popularity)=10),
Row(Director='Nicholson, Arch', Average=6.0, max(Popularity)=9, min(Popularity)=3),
Row(Director='Nigh, William', Average=71.0, max(Popularity)=71, min(Popularity)=71),
Row(Director='Nimoy, Leonard', Average=60.0, max(Popularity)=60, min(Popularity)=60),
Row(Director='Noyce, Phillip', Average=14.5, max(Popularity)=28, min(Popularity)=1),
Row(Director='Nuytten, Bruno', Average=32.0, max(Popularity)=32, min(Popularity)=32),
Row(Director='O'Connor, Pat', Average=57.0, max(Popularity)=57, min(Popularity)=57),
Row(Director='O'Herlihy, Michael', Average=62.0, max(Popularity)=62, min(Popularity)=62),
Row(Director='Oliansky, Joel', Average=45.0, max(Popularity)=45, min(Popularity)=45),
Row(Director='Oliver, Ron', Average=29.0, max(Popularity)=29, min(Popularity)=29),
Row(Director='Olmos, Edward James', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Ouellette, Jean-Paul', Average=36.0, max(Popularity)=36, min(Popularity)=36),
Row(Director='Owen, Cliff', Average=25.0, max(Popularity)=25, min(Popularity)=25),
Row(Director='Oz, Frank', Average=14.0, max(Popularity)=14, min(Popularity)=14),
Row(Director='Pabst, Georg Wilhelm', Average=73.0, max(Popularity)=73, min(Popularity)=73),
Row(Director='Pakula, Alan J.', Average=44.5, max(Popularity)=69, min(Popularity)=15),
Row(Director='Pal, George', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Palcy, Euzhan', Average=71.0, max(Popularity)=71, min(Popularity)=71),
Row(Director='Panama, Norman', Average=21.5, max(Popularity)=6, min(Popularity)=37),
Row(Director='Paris, Jerry', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Parker, Alan', Average=41.0, max(Popularity)=41, min(Popularity)=41),
Row(Director='Parrish, Robert', Average=80.0, max(Popularity)=80, min(Popularity)=80),
Row(Director='Patrick, John', Average=17.0, max(Popularity)=17, min(Popularity)=17),
Row(Director='Patterson, Richard', Average=59.0, max(Popularity)=59, min(Popularity)=59),
Row(Director='Patterson, Willi', Average=66.0, max(Popularity)=66, min(Popularity)=66),
Row(Director='Pearce, Richard', Average=11.0, max(Popularity)=11, min(Popularity)=11),
Row(Director='Peckinpah, Sam', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Peerce, Larry', Average=33.0, max(Popularity)=33, min(Popularity)=33),
Row(Director='Penn, Arthur', Average=12.0, max(Popularity)=12, min(Popularity)=12),
Row(Director='Penn, Leo', Average=13.0, max(Popularity)=13, min(Popularity)=13),
Row(Director='Perier, Etienne', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='Petrie, Daniel', Average=26.0, max(Popularity)=26, min(Popularity)=26),
Row(Director='Petrie, Donald', Average=74.0, max(Popularity)=74, min(Popularity)=74),
Row(Director='Pevney, Joseph', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Phillips, Lee', Average=36.0, max(Popularity)=36, min(Popularity)=36),
Row(Director='Phillips, Maurice', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Pialat, Maurice', Average=55.0, max(Popularity)=65, min(Popularity)=45),
Row(Director='Pillsbury, Sam', Average=80.0, max(Popularity)=80, min(Popularity)=80),
Row(Director='Plone, Allen', Average=85.0, max(Popularity)=85, min(Popularity)=85),
Row(Director='Poitier, Sidney', Average=38.0, max(Popularity)=40, min(Popularity)=36),
Row(Director='Polanski, Roman', Average=36.0, max(Popularity)=55, min(Popularity)=17),
Row(Director='Pollack, Sydney', Average=51.77777777777778, max(Popularity)=88, min(Popularity)=16),
Row(Director='Pontecorvo, Gillo', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Post, Ted', Average=54.0, max(Popularity)=67, min(Popularity)=28),
Row(Director='Preminger, Otto', Average=13.0, max(Popularity)=13, min(Popularity)=13),
Row(Director='Prince, Harold', Average=61.0, max(Popularity)=61, min(Popularity)=61),
Row(Director='Prior, David A.', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Puenzo, Luiz', Average=39.0, max(Popularity)=39, min(Popularity)=39),
Row(Director='Purcell, Evelyn', Average=52.0, max(Popularity)=52, min(Popularity)=52),
Row(Director='Pytko, Joe', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Pyun, Albert', Average=31.0, max(Popularity)=31, min(Popularity)=31),
Row(Director='Quine, Richard', Average=12.0, max(Popularity)=12, min(Popularity)=12),
Row(Director='Quinn, James', Average=64.0, max(Popularity)=64, min(Popularity)=64),
Row(Director='Quinn, John', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Quisenberry, Byron', Average=24.0, max(Popularity)=24, min(Popularity)=24),
Row(Director='Rademakers, Fons', Average=71.0, max(Popularity)=71, min(Popularity)=71),
Row(Director='Rafelson, Bob', Average=28.0, max(Popularity)=54, min(Popularity)=2),
Row(Director='Ramis, Harold', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='Rappeneau, Jean-Paul', Average=76.0, max(Popularity)=76, min(Popularity)=76),
Row(Director='Ray, Nicholas', Average=63.33333333333336, max(Popularity)=85, min(Popularity)=23),
Row(Director='Rayfield, David', Average=47.0, max(Popularity)=47, min(Popularity)=47),
Row(Director='Reed, Jerry', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Rees, Jerry', Average=84.0, max(Popularity)=84, min(Popularity)=84),
Row(Director='Reeve, Geoffrey', Average=42.0, max(Popularity)=42, min(Popularity)=42),
Row(Director='Reiner, Carl', Average=45.0, max(Popularity)=68, min(Popularity)=22),
Row(Director='Reiner, Rob', Average=48.0, max(Popularity)=48, min(Popularity)=48),
Row(Director='Reis, Irving', Average=25.0, max(Popularity)=25, min(Popularity)=25),
Row(Director='Reisz, Karel', Average=37.0, max(Popularity)=37, min(Popularity)=37),
Row(Director='Reitman, Ivan', Average=31.0, max(Popularity)=39, min(Popularity)=23),
Row(Director='Renoir, Jean', Average=33.0, max(Popularity)=33, min(Popularity)=33),
Row(Director='Resnais, Alain', Average=36.0, max(Popularity)=71, min(Popularity)=1),
Row(Director='Richards, Dick', Average=44.0, max(Popularity)=59, min(Popularity)=29),
Row(Director='Richardson, Tony', Average=32.5, max(Popularity)=39, min(Popularity)=26),
Row(Director='Rifkin, Adam', Average=77.0, max(Popularity)=77, min(Popularity)=77),
Row(Director='Ritchie, Michael', Average=53.33333333333336, max(Popularity)=86, min(Popularity)=22),
Row(Director='Ritt, Martin', Average=43.625, max(Popularity)=70, min(Popularity)=2),
Row(Director='Robertson, John S.', Average=73.0, max(Popularity)=73, min(Popularity)=73),
Row(Director='Robinson, John Mark', Average=20.0, max(Popularity)=20, min(Popularity)=20),
Row(Director='Rochlin, Sheldon', Average=82.0, max(Popularity)=82, min(Popularity)=82),
Row(Director='Roeg, Nicolas', Average=35.666666666666664, max(Popularity)=48, min(Popularity)=18),
Row(Director='Rohmer, Eric', Average=11.0, max(Popularity)=11, min(Popularity)=11),
Row(Director='Romer, George A.', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='Rosenberg, Stuart', Average=46.0, max(Popularity)=55, min(Popularity)=34),
Row(Director='Rosenthal, Rick', Average=47.0, max(Popularity)=87, min(Popularity)=7),
Row(Director='Ross, Herbert', Average=39.888888888888886, max(Popularity)=81, min(Popularity)=11),
Row(Director='Rossellini, Roberto', Average=57.25, max(Popularity)=69, min(Popularity)=34),
Row(Director='Rossen, Robert', Average=43.0, max(Popularity)=43, min(Popularity)=43),
Row(Director='Rossi, Franco', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Roth, Bobby', Average=71.0, max(Popularity)=71, min(Popularity)=71),
Row(Director='Ruben, Joseph', Average=63.0, max(Popularity)=63, min(Popularity)=63),
Row(Director='Ruben, Katt Shea', Average=80.0, max(Popularity)=80, min(Popularity)=80),
Row(Director='Rudolph, Alan', Average=57.0, max(Popularity)=57, min(Popularity)=57),
Row(Director='Ruggles, Wesley', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Rumar, Craig', Average=45.0, max(Popularity)=45, min(Popularity)=45),

```

```

Row(Director='Rush, Richard', Average=83.0, max(Popularity)=83, min(Popularity)=83),
Row(Director='Russell, Ken', Average=32.9, max(Popularity)=8, min(Popularity)=16),
Row(Director='Rydel, Mark', Average=40.5, max(Popularity)=58, min(Popularity)=23),
Row(Director='S, Jeremiah', Average=81.0, max(Popularity)=81, min(Popularity)=81),
Row(Director='Sagal, Boris', Average=80.0, max(Popularity)=80, min(Popularity)=80),
Row(Director='Saks, Gene', Average=53.5, max(Popularity)=67, min(Popularity)=40),
Row(Director='Sandrich, Jay', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Sanger, Jonathan', Average=22.0, max(Popularity)=22, min(Popularity)=22),
Row(Director='Sarafian, Richard C.', Average=40.0, max(Popularity)=40, min(Popularity)=40),
Row(Director='Sargent, Joseph', Average=54.0, max(Popularity)=54, min(Popularity)=54),
Row(Director='Savoca, Nancy', Average=66.0, max(Popularity)=66, min(Popularity)=66),
Row(Director='Sayles, John', Average=81.0, max(Popularity)=81, min(Popularity)=81),
Row(Director='Schaffner, Franklin J.', Average=8.5, max(Popularity)=9, min(Popularity)=8),
Row(Director='Schatzberg, Jerry', Average=10.0, max(Popularity)=10, min(Popularity)=10),
Row(Director='Scheerer, Robert', Average=3.0, max(Popularity)=3, min(Popularity)=3),
Row(Director='Schepisi, Fred', Average=36.25, max(Popularity)=9, min(Popularity)=3),
Row(Director='Schlesinger, John', Average=54.75, max(Popularity)=81, min(Popularity)=33),
Row(Director='Schl ndorff, Volker', Average=13.0, max(Popularity)=13, min(Popularity)=13),
Row(Director='Schmoeller, David', Average=20.0, max(Popularity)=20, min(Popularity)=20),
Row(Director='Schnitzer, Robert Allen', Average=41.0, max(Popularity)=41, min(Popularity)=41),
Row(Director='Schrader, Paul', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Schroeder, Barbet', Average=45.0, max(Popularity)=73, min(Popularity)=23),
Row(Director='Schultz, Michael', Average=69.0, max(Popularity)=69, min(Popularity)=69),
Row(Director='Schumacher, Joel', Average=43.0, max(Popularity)=67, min(Popularity)=19),
Row(Director='Scola, Ettore', Average=80.0, max(Popularity)=80, min(Popularity)=80),
Row(Director='Scorsese, Martin', Average=32.25, max(Popularity)=84, min(Popularity)=15),
Row(Director='Scott, Ridley', Average=49.75, max(Popularity)=83, min(Popularity)=1),
Row(Director='Scott, Tony', Average=16.0, max(Popularity)=8, min(Popularity)=3),
Row(Director='Seaton, George', Average=0.0, max(Popularity)=0, min(Popularity)=0),
Row(Director='Sebastian, Beverly', Average=50.5, max(Popularity)=73, min(Popularity)=28),
Row(Director='Segal, Alex', Average=73.0, max(Popularity)=73, min(Popularity)=73),
Row(Director='Seidelman, Susan', Average=42.33333333333336, max(Popularity)=43, min(Popularity)=41),
Row(Director='Seiter, William A.', Average=70.0, max(Popularity)=87, min(Popularity)=53),
Row(Director='Seltzer, David', Average=11.0, max(Popularity)=11, min(Popularity)=11),
Row(Director='Serious, Yahoo', Average=47.0, max(Popularity)=47, min(Popularity)=47),
Row(Director='Sevush, Herb', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Sharman, Jim', Average=59.0, max(Popularity)=59, min(Popularity)=59),
Row(Director='Shatner, William ', Average=87.0, max(Popularity)=87, min(Popularity)=87),
Row(Director='Sheridan, Jim', Average=32.0, max(Popularity)=32, min(Popularity)=32),
Row(Director='Sherman, Gary', Average=28.0, max(Popularity)=28, min(Popularity)=28),
Row(Director='Sherman, George', Average=61.33333333333336, max(Popularity)=83, min(Popularity)=33),
Row(Director='Sherman, Vincent', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Sholder, Jack', Average=75.0, max(Popularity)=75, min(Popularity)=75),
Row(Director='Sidaris, Andy', Average=29.0, max(Popularity)=36, min(Popularity)=20),
Row(Director='Siegel, Don', Average=49.4, max(Popularity)=72, min(Popularity)=22),
Row(Director='Silver, Joan Micklin', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Silver, Marisa', Average=42.0, max(Popularity)=42, min(Popularity)=42),
Row(Director='Silverstein, Elliot', Average=62.0, max(Popularity)=62, min(Popularity)=62),
Row(Director='Simon, S. Sylvan', Average=57.0, max(Popularity)=57, min(Popularity)=57),
Row(Director='Sinise, Gary', Average=53.0, max(Popularity)=53, min(Popularity)=53),
Row(Director='Siodmak, Robert', Average=77.0, max(Popularity)=77, min(Popularity)=77),
Row(Director='Smight, Jack', Average=36.0, max(Popularity)=36, min(Popularity)=36),
Row(Director='Smith, Charles Martin', Average=47.0, max(Popularity)=47, min(Popularity)=47),
Row(Director='Smithey, Alan', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Soderbergh, Steven', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='Sonnenfeld, B.', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Spielberg, Steven', Average=17.1, max(Popularity)=8, min(Popularity)=24),
Row(Director='Spottiswoode, Roger', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Stanley, Paul', Average=27.0, max(Popularity)=27, min(Popularity)=27),
Row(Director='Stein, Ken', Average=10.0, max(Popularity)=10, min(Popularity)=10),
Row(Director='Stern, Sander', Average=47.0, max(Popularity)=88, min(Popularity)=6),
Row(Director='Stern, Steven Hilliard', Average=81.5, max(Popularity)=88, min(Popularity)=75),
Row(Director='Sternberg, Josef von', Average=43.0, max(Popularity)=43, min(Popularity)=43),
Row(Director='Stevens, George', Average=43.5, max(Popularity)=61, min(Popularity)=26),
Row(Director='Stevens, George, Jr.', Average=56.0, max(Popularity)=56, min(Popularity)=56),
Row(Director='Stevenson, Robert', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Stiller, Mauritz', Average=63.0, max(Popularity)=63, min(Popularity)=63),
Row(Director='Stockwell, John', Average=36.0, max(Popularity)=36, min(Popularity)=36),
Row(Director='Stone, Oliver', Average=42.5, max(Popularity)=8, min(Popularity)=44),
Row(Director='Streisand, Barbra', Average=46.0, max(Popularity)=46, min(Popularity)=46),
Row(Director='Strock, Herbert L.', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Stuart, Mel', Average=65.0, max(Popularity)=65, min(Popularity)=65),
Row(Director='Sturges, John', Average=78.6666666666667, max(Popularity)=84, min(Popularity)=73),
Row(Director='Surin, Fred', Average=70.0, max(Popularity)=70, min(Popularity)=70),
Row(Director='Szab , Istv n', Average=77.0, max(Popularity)=80, min(Popularity)=74),
Row(Director='Szwarc, Jeannot', Average=19.0, max(Popularity)=19, min(Popularity)=19),
Row(Director='Tatoulis, Colin South, John', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Taurog, Norman', Average=21.0, max(Popularity)=21, min(Popularity)=21),
Row(Director='Taylor, Baz', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Taylor, Dan', Average=10.0, max(Popularity)=10, min(Popularity)=10),
Row(Director='Taylor, Don', Average=39.5, max(Popularity)=58, min(Popularity)=11),
Row(Director='Taylor, Gavin', Average=48.0, max(Popularity)=48, min(Popularity)=48),
Row(Director='Taylor, Jud', Average=52.0, max(Popularity)=86, min(Popularity)=2),
Row(Director='Taylor, Ray', Average=54.75, max(Popularity)=87, min(Popularity)=17),
Row(Director='Taylor, Roderick', Average=47.0, max(Popularity)=47, min(Popularity)=47),
Row(Director='Taylor, Sam', Average=43.33333333333336, max(Popularity)=66, min(Popularity)=31),
Row(Director='Teague, Lewis', Average=68.0, max(Popularity)=68, min(Popularity)=68),
Row(Director='Theumer, Ernst R. von', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Thomas, Dave', Average=67.0, max(Popularity)=67, min(Popularity)=67),
Row(Director='Thomas, Gerald', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Thomas, Ralph', Average=83.0, max(Popularity)=83, min(Popularity)=83),
Row(Director='Thompson, Brett', Average=19.0, max(Popularity)=19, min(Popularity)=19),
Row(Director='Thompson, J. Lee', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Thorpe, Richard', Average=11.0, max(Popularity)=11, min(Popularity)=11),
Row(Director='Totten, Robert', Average=86.0, max(Popularity)=86, min(Popularity)=86),
Row(Director='Tourneur, Jacques', Average=7.0, max(Popularity)=7, min(Popularity)=7),
Row(Director='Towne, Robert', Average=50.0, max(Popularity)=50, min(Popularity)=50),
Row(Director='Towns, Philip Saville', Average=39.0, max(Popularity)=39, min(Popularity)=39),
Row(Director='Travis, Mark W.', Average=30.0, max(Popularity)=30, min(Popularity)=30),

```

```
Row(Director='Trikonis, Gus', Average=62.0, max(Popularity)=62, min(Popularity)=62),
Row(Director='Troell, Jan', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Truffaut, François', Average=52.666666666666664, max(Popularity)=82, min(Popularity)=10),
Row(Director='Tuchner, Michael', Average=34.5, max(Popularity)=36, min(Popularity)=33),
Row(Director='Vadim, Roger', Average=25.333333333333332, max(Popularity)=8, min(Popularity)=29),
Row(Director='Van Peebles, Mario', Average=80.0, max(Popularity)=80, min(Popularity)=80),
Row(Director='Vanwagenen, Sterling', Average=3.0, max(Popularity)=3, min(Popularity)=3),
Row(Director='Varda, Agnes', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Veber, Francis', Average=39.0, max(Popularity)=54, min(Popularity)=24),
Row(Director='Verhoeven, Paul', Average=24.5, max(Popularity)=8, min(Popularity)=41),
Row(Director='Vidor, Charles', Average=54.0, max(Popularity)=57, min(Popularity)=48),
Row(Director='Vidor, King', Average=51.0, max(Popularity)=51, min(Popularity)=51),
Row(Director='Vigne, Daniel', Average=57.5, max(Popularity)=64, min(Popularity)=51),
Row(Director='Vorhaus, Bernard', Average=38.0, max(Popularity)=38, min(Popularity)=38),
Row(Director='Waggner, George', Average=74.0, max(Popularity)=74, min(Popularity)=74),
Row(Director='Wajda, Andrzej', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Walsh, Raoul', Average=37.0, max(Popularity)=52, min(Popularity)=22),
Row(Director='Walters, Charles', Average=67.0, max(Popularity)=67, min(Popularity)=67),
Row(Director='Walton, Kip', Average=33.5, max(Popularity)=54, min(Popularity)=13),
Row(Director='Ward, David S.', Average=64.0, max(Popularity)=64, min(Popularity)=64),
Row(Director='Waris, Hussein', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Wayne, John', Average=29.0, max(Popularity)=29, min(Popularity)=29),
Row(Director='Webb, Millard', Average=None, max(Popularity)=None, min(Popularity)=None),
Row(Director='Weeks, Stephen', Average=5.0, max(Popularity)=5, min(Popularity)=5),
Row(Director='Weill, Claudia', Average=0.0, max(Popularity)=0, min(Popularity)=0),
Row(Director='Weir, Peter', Average=30.666666666666668, max(Popularity)=8, min(Popularity)=25),
Row(Director='Weis, Gary', Average=25.0, max(Popularity)=25, min(Popularity)=25),
Row(Director='Welles, Orson', Average=39.666666666666664, max(Popularity)=80, min(Popularity)=16),
Row(Director='Wellman, William', Average=17.5, max(Popularity)=7, min(Popularity)=15),
Row(Director='Wenders, Wim', Average=52.625, max(Popularity)=73, min(Popularity)=27),
Row(Director='Wendkos, Paul', Average=69.0, max(Popularity)=69, min(Popularity)=69),
Row(Director='Werner, Peter', Average=17.0, max(Popularity)=17, min(Popularity)=17),
Row(Director='Wertmuller, Lina', Average=52.0, max(Popularity)=52, min(Popularity)=52),
Row(Director='Wharmby, Tony', Average=41.0, max(Popularity)=41, min(Popularity)=41),
Row(Director='Wicki, Bernhard', Average=9.0, max(Popularity)=9, min(Popularity)=9),
Row(Director='Widerberg, Bo', Average=28.0, max(Popularity)=28, min(Popularity)=28),
Row(Director='Wiemer, Robert', Average=43.0, max(Popularity)=43, min(Popularity)=43),
Row(Director='Wilcox, Fred M', Average=79.0, max(Popularity)=79, min(Popularity)=79),
Row(Director='Wilder, Billy', Average=3.0, max(Popularity)=3, min(Popularity)=3),
Row(Director='Wilder, Gene', Average=42.0, max(Popularity)=42, min(Popularity)=42),
Row(Director='Wilmington, Michael', Average=10.0, max(Popularity)=10, min(Popularity)=10),
Row(Director='Wilson, Hugh', Average=44.0, max(Popularity)=44, min(Popularity)=44),
Row(Director='Wimhurst, Jolyon', Average=85.0, max(Popularity)=85, min(Popularity)=85),
Row(Director='Winkler, Henry', Average=23.0, max(Popularity)=23, min(Popularity)=23),
Row(Director='Winkler, Irwin', Average=88.0, max(Popularity)=88, min(Popularity)=88),
Row(Director='Winner, Michael', Average=4.0, max(Popularity)=8, min(Popularity)=0),
Row(Director='Winning, David', Average=34.5, max(Popularity)=8, min(Popularity)=61),
Row(Director='Wise, Robert', Average=57.5, max(Popularity)=77, min(Popularity)=38),
Row(Director='Wolfe, Donald', Average=24.0, max(Popularity)=24, min(Popularity)=24),
Row(Director='Wrede, Caspar', Average=4.0, max(Popularity)=4, min(Popularity)=4),
Row(Director='Wright, Thomas J.', Average=35.0, max(Popularity)=35, min(Popularity)=35),
Row(Director='Wyler, William', Average=57.0, max(Popularity)=81, min(Popularity)=30),
Row(Director='Yorkin, Bud', Average=62.0, max(Popularity)=62, min(Popularity)=62),
Row(Director='Young, Robert M.', Average=49.0, max(Popularity)=49, min(Popularity)=49),
Row(Director='Young, Terence', Average=7.0, max(Popularity)=7, min(Popularity)=7),
Row(Director='Young, Terrence', Average=8.0, max(Popularity)=8, min(Popularity)=8),
Row(Director='Zanussi, Krzystoff', Average=81.0, max(Popularity)=84, min(Popularity)=78),
Row(Director='Zeffirelli, Franco', Average=20.0, max(Popularity)=20, min(Popularity)=20),
Row(Director='Zemeckis, Robert', Average=35.0, max(Popularity)=9, min(Popularity)=61),
Row(Director='Zemeckis, Robert ', Average=83.0, max(Popularity)=83, min(Popularity)=83),
Row(Director='Zieff, Howard', Average=67.0, max(Popularity)=73, min(Popularity)=61),
Row(Director='Zinnemann, Fred', Average=40.666666666666664, max(Popularity)=75, min(Popularity)=20),
Row(Director='Zucker, David', Average=9.0, max(Popularity)=9, min(Popularity)=9),
Row(Director='Zucker, Jerry', Average=6.0, max(Popularity)=6, min(Popularity)=6),
Row(Director='Zwick, Edward', Average=66.0, max(Popularity)=66, min(Popularity)=66)]
```

Question 2 In this example, you will use the given "Movies.csv" data file. Your task is to write an efficient spark code for finding movies (released after 1970) that might be similar or related to a movie made in the 1960's (1960-1969). To keep things simple, for comparing two movies we will use the following fields: Genre, director, lead actor, lead actress. So, we can say two movies might be similar if they have same genre, director and same lead actor or actress.

```
In [ ]: df = spark.read.option("delimiter", ";").option("header", "true").option("encoding", "ISO-8859-1").csv('Movies.csv')
A=df.filter(col('Year')>1970)
B=df.filter((col('Year')>=1960) & (col('Year')<=1969))
C=A.join(B,['Genre', "Director", "Actor", "Actress"])
C.show()
```

Genre	Director	Actor	Actress	Year	Length	Title	Popularity	Awards	*Image,,,,,	Year	Length	Title	Popularity
Drama	LeLouch, Claude	Trintignant, Jean...	Aimee, Anouk	1986	112	A Man & a Woman: ...	49	No	NicholasCage.png,,	1966	103	Un Hombre y una M...	6
									Yes	NicholasCage.png,,			
Drama	LeLouch, Claude	Trintignant, Jean...	Aimee, Anouk	1986	112	A Man & a Woman: ...	49	No	NicholasCage.png,,	1966	103	A Man & a Woman	46
									Yes	NicholasCage.png,,			

```
In [ ]: # Stop the Spark session when you're done
spark.stop()
```