DATABASE MODELLING AND DATA ANALYTICS

DATA ANALYTICS REPORT

ON

WAR SIMULATION GAME

Pritisha Mishra
s5331747@Bournemouth.ac.uk

Nisha Pal
S5324272@Bournemouth.ac.uk

Vivek Dilipbhai Mistry
S5425981@Bournemouth.ac.uk

Table Of Contents:

1. Task 1:

1.1 ENTITY RELATIONSHIP DIAGRAM(ERD): Bike Rental

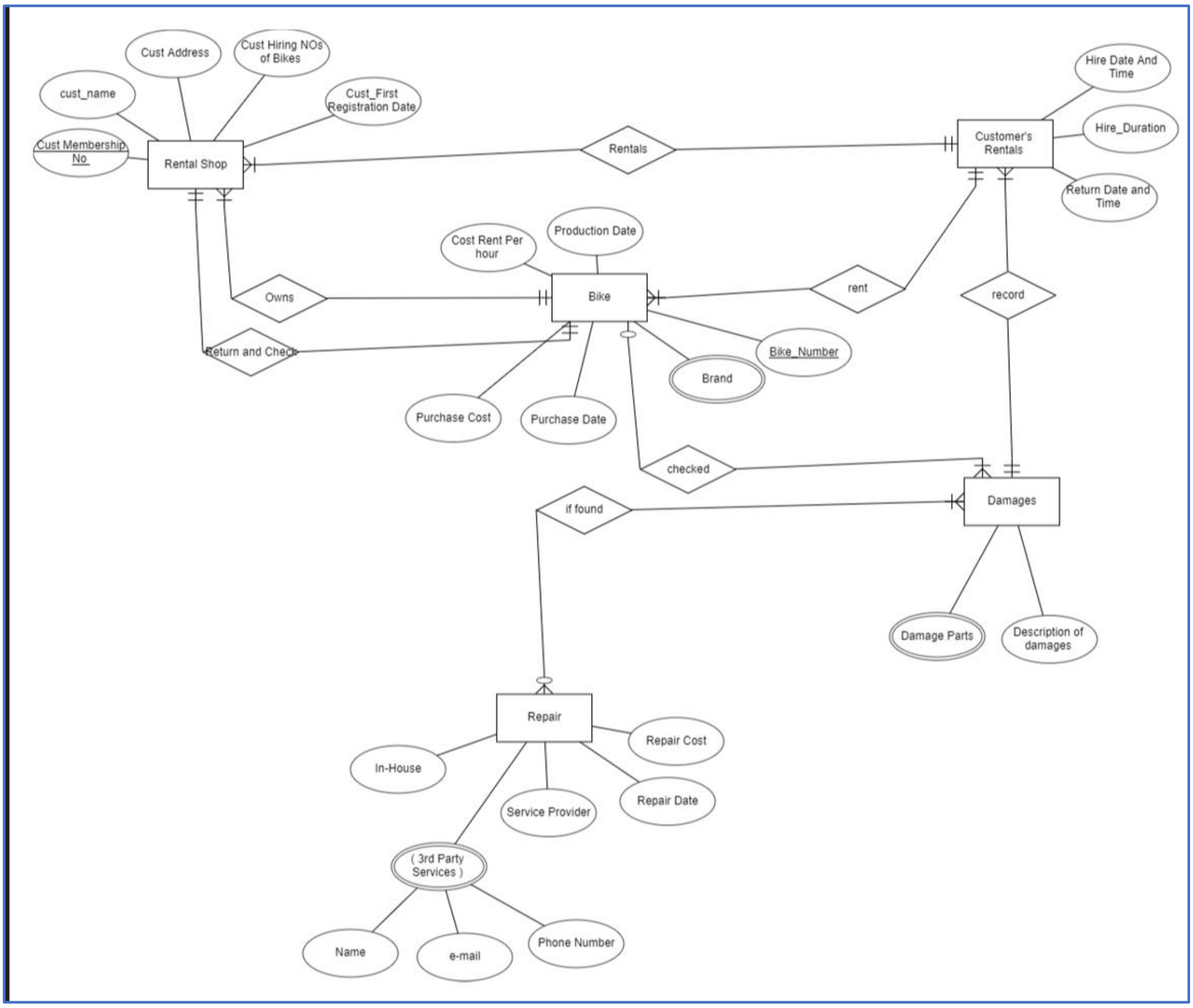


Diagram 1: ERD Diagram

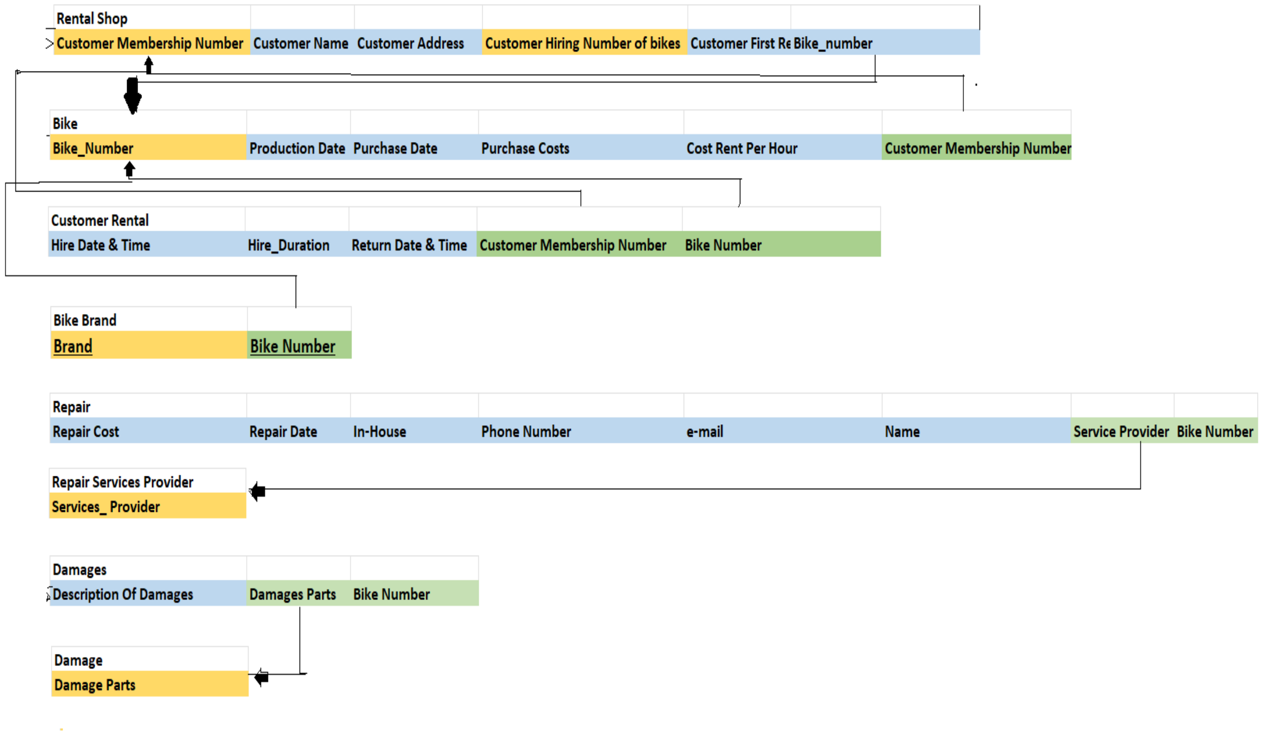
ASSUMPTIONS MADE:

- Customers hire the bike and register to the shop via online or login process.
- After Hiring the bike payment and method of payment should be declared and invoice should be generated with unique invoice number, hence it is helpful for both shop and customer .
- After Return the bike it should be recorded from Bike Number with the invoice number.

CARIDINALITY EXPLANATIONS:

- Rental shop is mandatory for customer rental and for one shop there are numerous rentals.
- Customer rental is recorded if there is any damage, it is mandatory and it should be many.
- If damage is found it need to repair by service provider or 3rd party services .There are many 3rd party services so it should mention their name ,email and phone number.
- It is mandatory to check each bike after return and if damages found it need to be repaired and recorded.
- Bike have multiple brands , damage of the bike have multiple parts and there are multiple service providers for repairing the damages.
- Rental shop have mandatory bikes with different brands and bike number, shop has all the details of customer with their registration number.

1.2 RELATIONAL DIAGRAM:



Rental Shop

Customer Membership Number	Customer Name	Customer Address	Customer Hiring Number of bikes	Customer First Re	Bike_number

Bike

Bike_Number	Production Date	Purchase Date	Purchase Costs	Cost Rent Per Hour	Customer Membership Number

Customer Rental

Hire Date & Time	Hire_Duration	Return Date & Time	Customer Membership Number	Bike Number

Bike Brand

Brand	Bike Number

Repair

Repair Cost	Repair Date	In-House	Phone Number	e-mail	Name	Service Provider	Bike Number

Repair Services Provider

Services_ Provider

Damages

Description Of Damages	Damages Parts	Bike Number

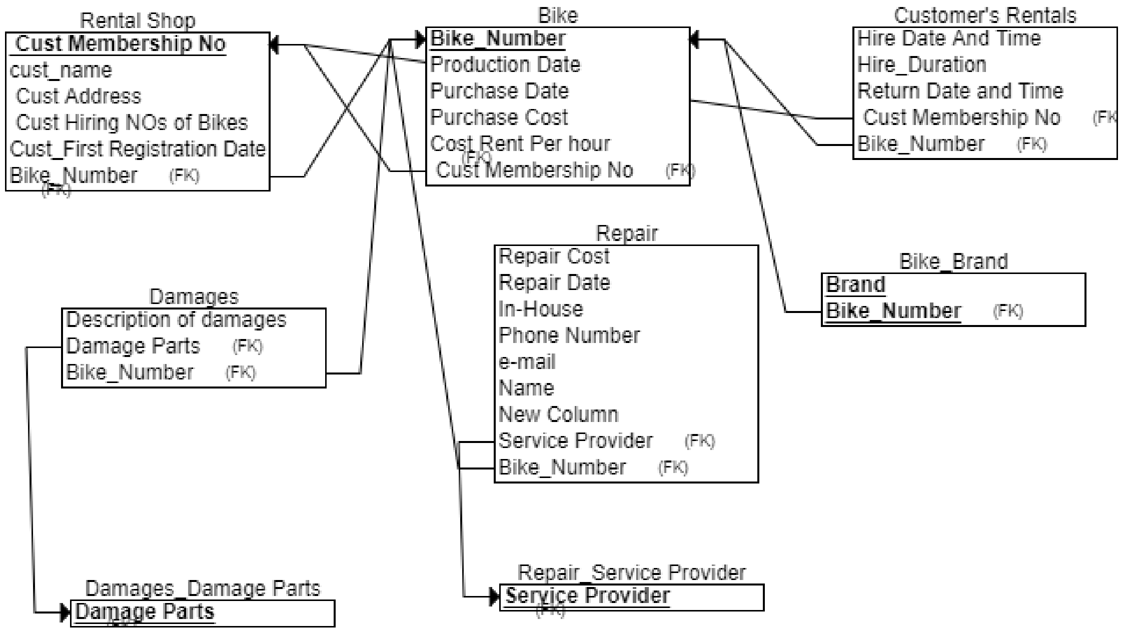
Damage

Damage Parts

Diagram 2: Relational Diagram

Notes:

The background yellow denotes the primary Key, the light green box background denotes Foreign Key.



2. TASK 2:

KAGGLE PUBLIC LEADERBOARD SCORE: 2.06173

PUBLIC LEADERBOARD : RANK:2

INTRODUCTION

2.1 Business Understanding

2.1.1 Business Objective:

The WAR Game's goal is to strengthen corporate and planning activities while also understanding the business strategy of the War Game. The actual purpose of War games is strategy games that cover the most serious conflicts, so the developer's primary goal is to analyse the game and reduce the complexity of the game. William Uricchio's chapter 'Simulation, History, and Computer Games,' published in the Handbook of Computer Game Studies (Raessens and Goldstein 2005), went beyond a single focus on the content of individual historical games in isolation or its application to other domains (such as education). The study of games that are linked to historical debates or depict the past in some way' (Chapman 2016, 16). Squire's (2004) thesis on the use of Sid Meier's Civilization (1991–2016) to teach history in schools, for example, demonstrates a mastery of historical game studies as well as the broader subject of games and learning.

"Another goal of the research, as well as developer opinions on the relationship between history and games5, is to examine the relationship between cultural/collective memory and games." (Begy 2015; Pötzsch and Šisler 2016);

The actual purpose and scope of the game, the team represents distinct rivals.

War games should incorporate specific real-world scenarios to be most effective. One of these is a very competitive market where players must react to one another's actions. Unpredictability is another factor, as seen by shifting technologies and market demand. A long-term view is also necessary to demonstrate how current decisions may affect future profitability. Managers understand the significance of being very explicit in their communications with the market, which is one of the most essential outcomes of wargames.

Despite the fact that each simulation is unique, there are a few things that a corporation may anticipate to gain from the experience. The first is that your perspective of the world will transform, as if you were rotating a map over to see what was on the other side of the mountain. The assumptions about organisation, competitors, and industry will have been tested implicitly and explicitly. Some assumptions will have survived, while others will have been rejected, resulting in a new perspective. And because of this new perspective, topics that would have never crossed the mind before will suddenly appear as opportunities, some old bright ideas will be put to rest, and this will be significantly more productive the traps that lurk everywhere.

2.1.2 Project Plan:

Using the supplied Historical Military Battles Dataset (Kaggle,2021) and as well as the CRISP-DM technique(Chapman et al 2000), this paper will develop a predictive model which can predict the relative achievement of the attacker, measured via the difference of attackers and defenders achievement values, as well as to understand which factors contribute to determining the achievement.

Throughout this project, the Oracle SQL Developer (18.1) is used to modify the supplied database with the purpose of creating datasets that may be used in algorithms to increase prediction accuracy. Microsoft (2021) Excel is used to visualize the data through many charts to enhance understanding of data and correlation. To develop prediction models, the data science application weak is employed (University of Waikato 2020). An online model checker will score the forecasts (Kaggle).

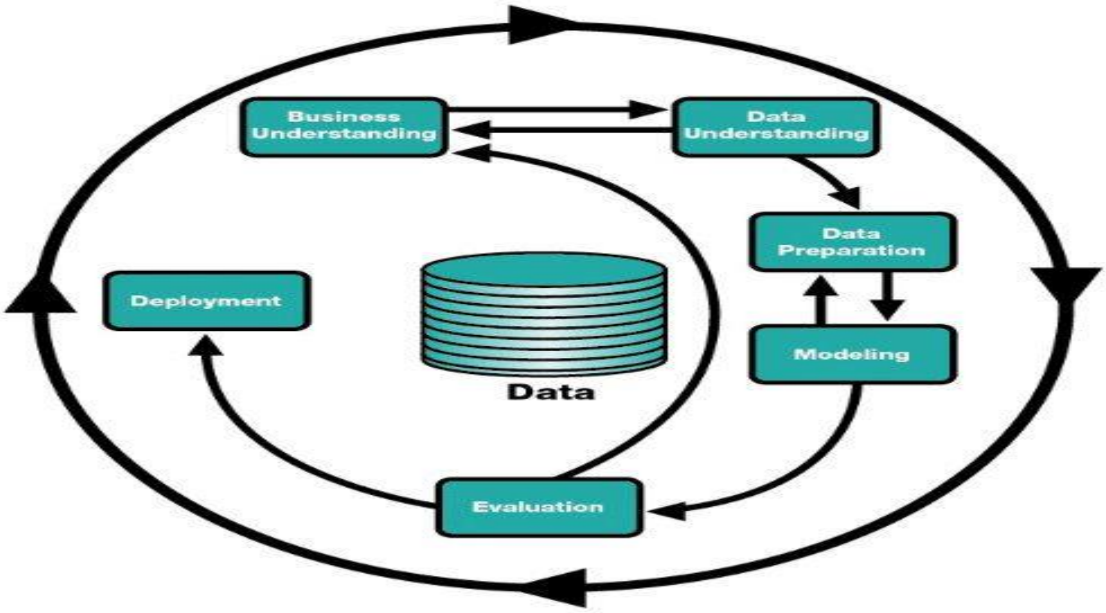


Fig1: CRISP- DM METHOD

Situation Assessment:

The predictive model will predict the relative achievement of the attacker. Extra planning is designed to prevent risks with the model.

- When cleaning a dataset for data mining, one concern is to delete incorrect data, missing values affects the performance of the algorithms, so need to be replaced. Most of the data mining algorithms will benefit by having only a single value per attribute for a data instance, so multiple values per attribute need to be transferred.

- Another potential risk is underfitting and overfitting when the predictive model describes occasional error or noise rather than underlying relationship and Relations in data are more complex than the chosen model. This can be avoided if more appropriate data is used during the training process. Multiple solutions, such as algorithm-specific procedures (regularisation) and cross-validation, or reserving unseen data for testing, can be employed to mitigate the hazards. This restricts the number of modifications that can be made (Frank et al 2011).

DATA UNDERSTANING- Historical Military Battles Database:

The data used in this report originates from the Historical Military Battles Dataset(Kaggle,2021) and is hosted on the Bournemouth University's Oracle server. To access the data using the Oracle SQL Developer. A remote connection to foston.bournemouth.ac.uk:1948 needs to be established.

The remote Historical Military database contains 6-tables.Table 1 represents the main properties of each table .To access and visualise each table the view is created in SQL separately for every table and exported in csv for use in Weka.

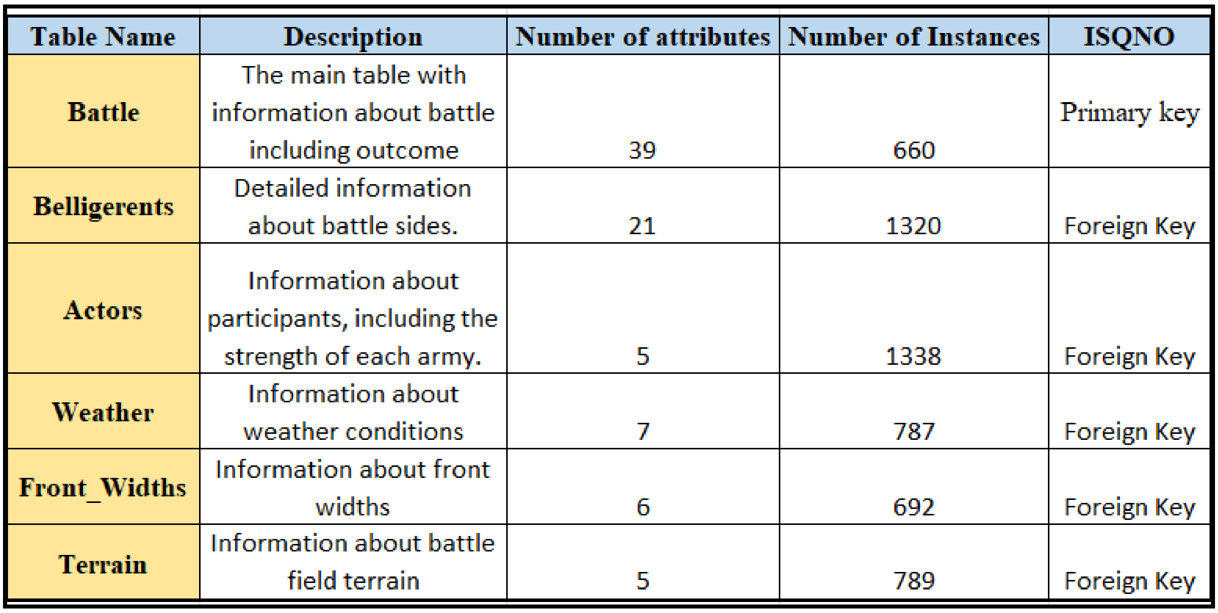


Table Name	Description	Number of attributes	Number of Instances	ISQNO
Battle	The main table with information about battle including outcome	39	660	Primary key
Belligerents	Detailed information about battle sides.	21	1320	Foreign Key
Actors	Information about participants, including the strength of each army.	5	1338	Foreign Key
Weather	Information about weather conditions	7	787	Foreign Key
Front_Widths	Information about front widths	6	692	Foreign Key
Terrain	Information about battle field terrain	5	789	Foreign Key

Table 1:Properties of each table

BATTLE TABLE:

This is the parent table with all information about the Battle including outcome as WINA attribute. There are 132(20%) missing values for Wina attribute in the Battle table. War initiator gives information about the attacker.

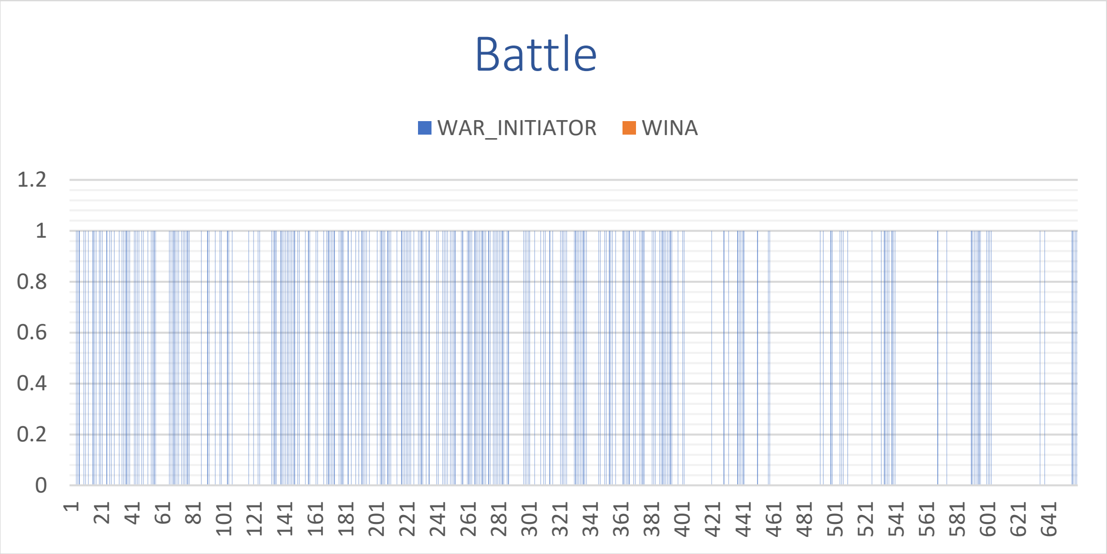
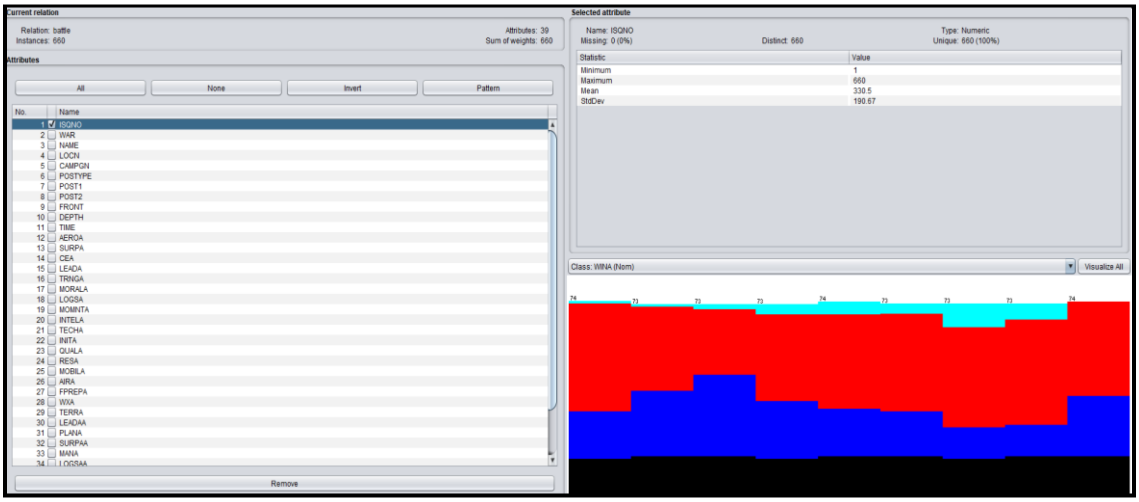


Chart1: War Initiator, Wina attribute

The SQL query used to create the visualization is :

```
CREATE VIEW BATTLE_TEST AS
SELECT* FROM BATTLE;
```



Relation: battle
Instances: 660

Attributes: 39
Sum of weights: 660

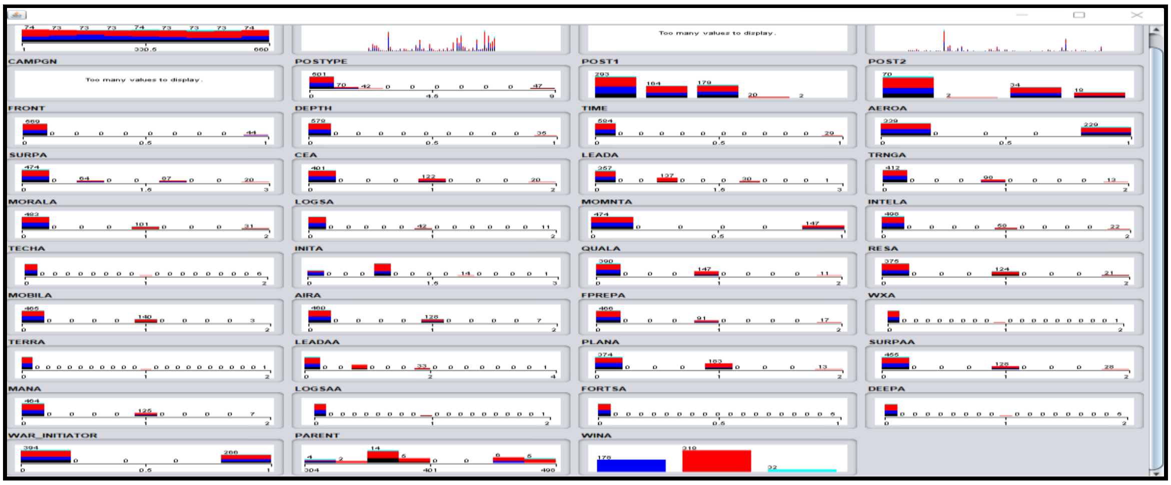
No.		Name
1	☑	ISQNO
2	□	WAR
3	□	NAME
4	□	LOCN
5	□	CAMPGN
6	□	POSTYPE
7	□	POST1
8	□	POST2
9	□	FRONT
10	□	DEPTH
11	□	TIME
12	□	AEROA
13	□	SURPA
14	□	CEA
15	□	LEADA
16	□	TRNGA
17	□	MORALA
18	□	LOGSA
19	□	MOMNTA
20	□	INTELA
21	□	TECHA
22	□	INITA
23	□	QUALA
24	□	RESA
25	□	MOBILA
26	□	AIRA
27	□	FPREPA
28	□	WXA
29	□	TERRA
30	□	LEADAA
31	□	PLANA
32	□	SURPAA
33	□	MANA
34	□	LOGSAA

Name: ISQNO
Missing: 0 (0%)
Distinct: 660
Type: Numeric
Unique: 660 (100%)

Statistic	Value
Minimum	1
Maximum	660
Mean	330.5
StdDev	190.67

Class: WINA (Nom)

Picture1: Attributes in Battle



Picture2: Visualization of Attributes Battle

BELLIGERENTS TABLE:

This table represents the detail information about Battle sides. Each Battle have two Belligerents. We also need to use the ach attribute from Belligerents, which determines the result of the battle. The ACH attribute is having 264(20%) null values. When Attacker column is 0, the corresponding belligerent is defender, when this is 1, the corresponding belligerent is attacker.

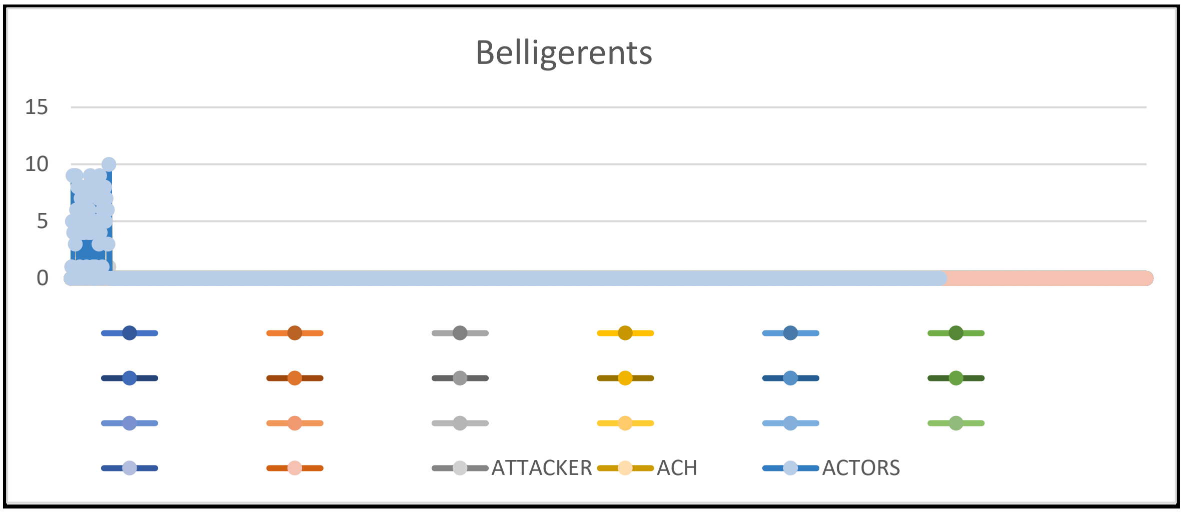
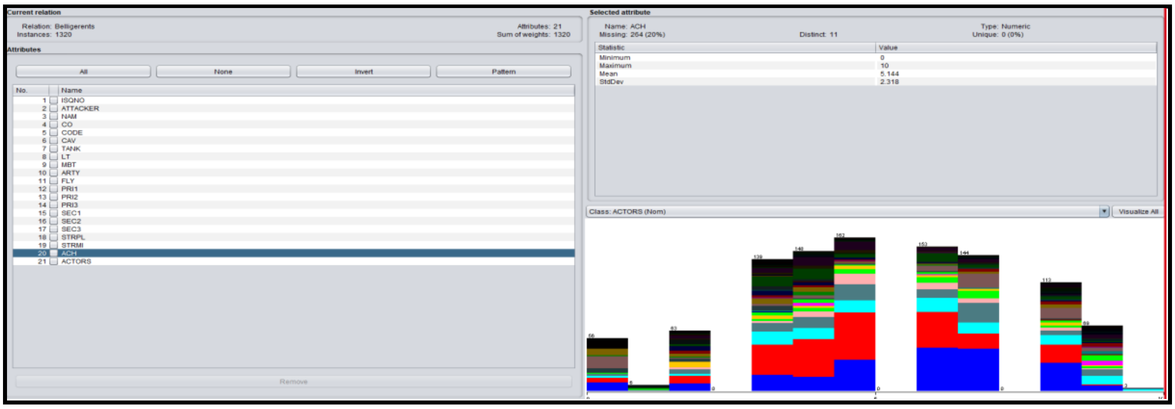


Chart2: Attacker,Ach,Actors

The SQL query used to create the visualization is :

```
CREATE VIEW BELLIGERENTS_TEST AS
SELECT* FROM BELLIGERENTS;
```



Current relation

Relation: Belligerents
Instances: 1320

Attributes: 21
Sum of weights: 1320

Attributes

All | None | Invert | Pattern

No.		Name
1	□	ISQNO
2	□	ATTACKER
3	□	NAM
4	□	CO
5	□	CODE
6	□	CAV
7	□	TANK
8	□	LT
9	□	MBT
10	□	ARTY
11	□	FLY
12	□	PRI1
13	□	PRI2
14	□	PRI3
15	□	SEC1
16	□	SEC2
17	□	SEC3
18	□	STRPL
19	□	STRMI
20	□	ACH
21	□	ACTORS

Remove

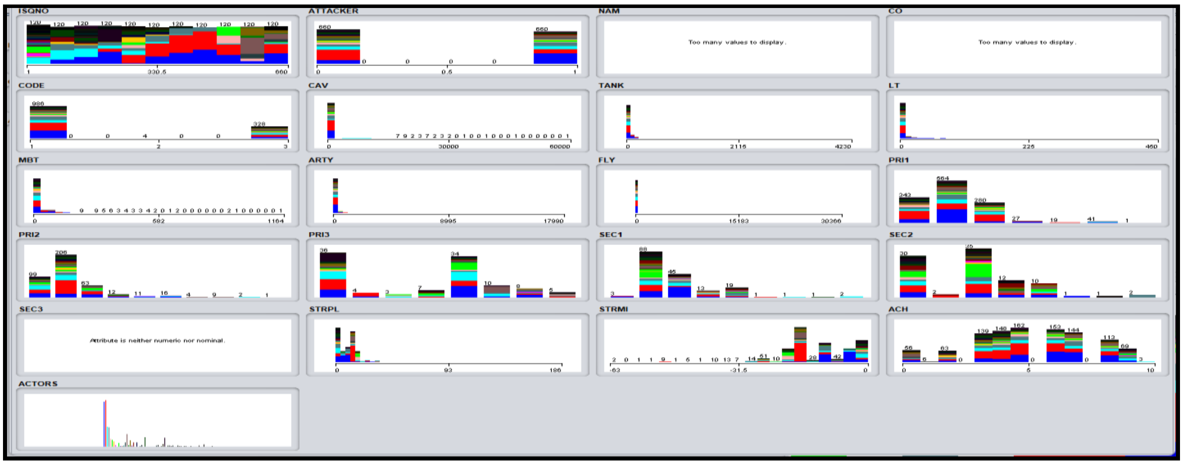
Selected attribute

Name: ACH
Missing: 264 (20%)
Distinct: 11
Type: Numeric
Unique: 0 (0%)

Statistic	Value
Minimum	0
Maximum	10
Mean	5.144
StdDev	2.318

Class: ACTORS (Nom)

Picture3: Attributes in Belligerent



Too many values to display.

Too many values to display.

Attribute is neither numeric nor nominal.

Picture4: Visualization of Attributes Belligerents

WEATHER TABLE:

This is the Precipitation, Temperature, Cloudiness, Season and cloud weather conditions WXNo as the weather condition 1 declared at the start of the Battle or 2 and 3 which is at the end of the Battle.

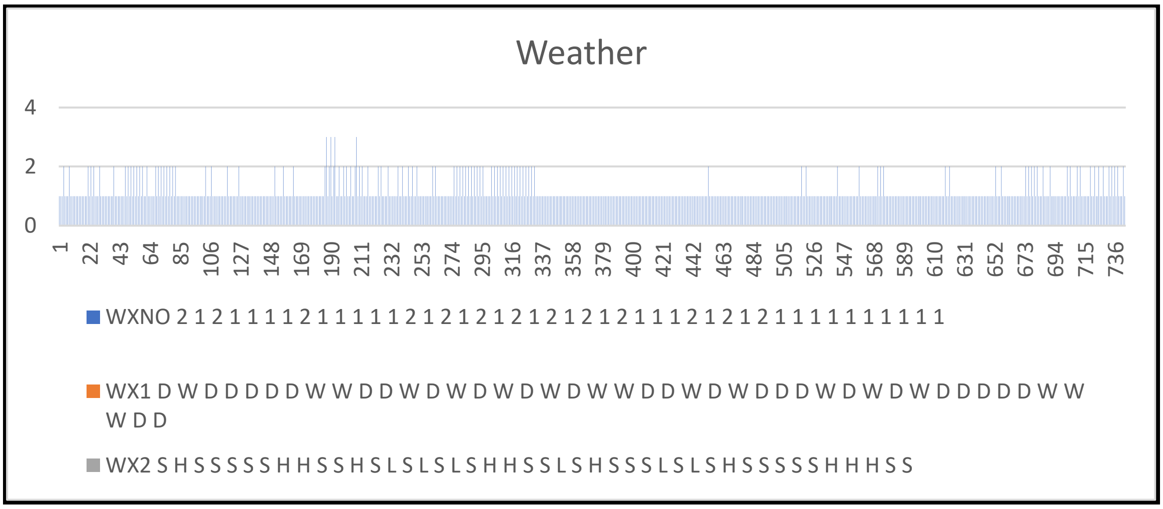
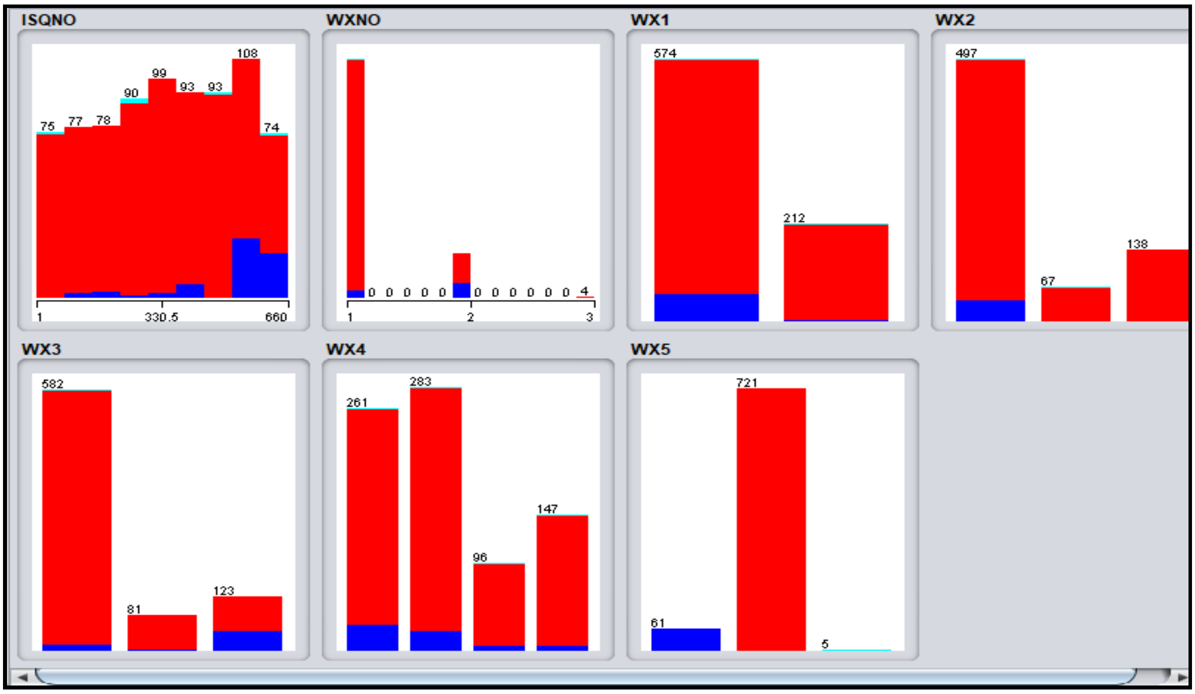


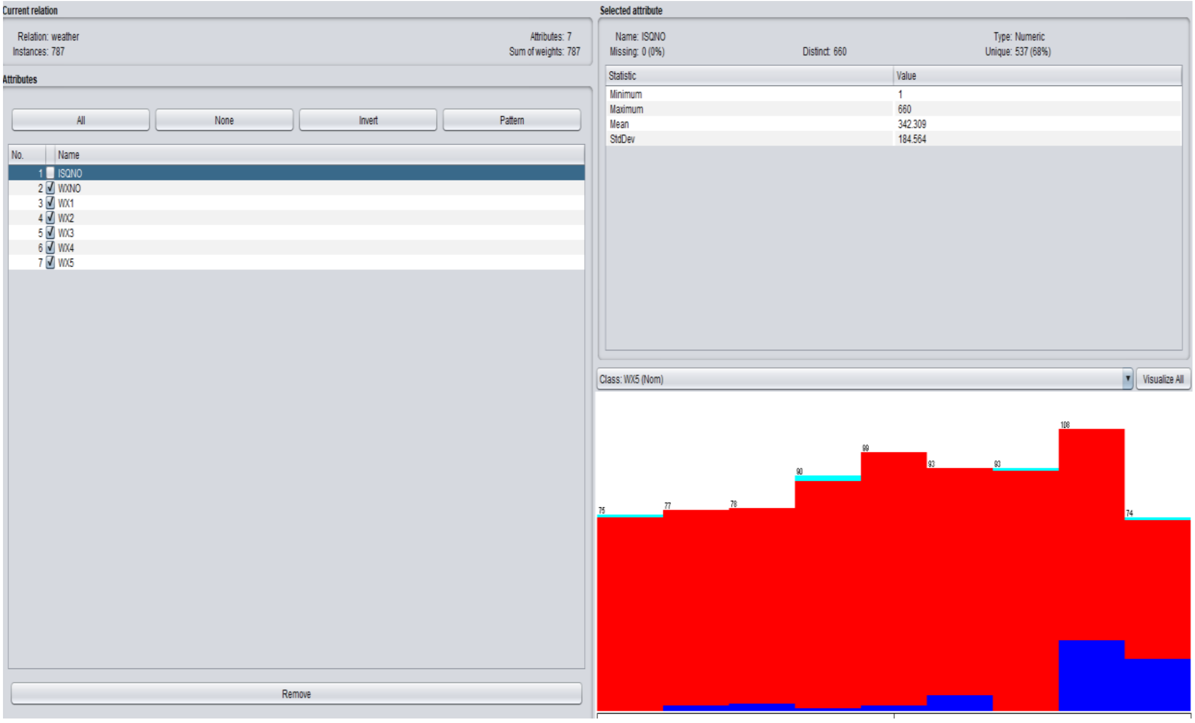
Chart3: WXno,WX1,WX2

The SQL query used to create the visualization is :

```
CREATE VIEW WEATHER_TEST AS
SELECT* FROM WEATHER;
```



Picture5: Attributes in Weather



Current relation

Relation: weather
Instances: 787

Attributes: 7
Sum of weights: 787

Attributes

No.		Name
1	□	ISQNO
2	☑	WXNO
3	☑	WX1
4	☑	WX2
5	☑	WX3
6	☑	WX4
7	☑	WX5

Selected attribute

Name: ISQNO
Missing: 0 (0%)
Distinct: 660
Type: Numeric
Unique: 537 (68%)

Statistic	Value
Minimum	1
Maximum	660
Mean	342.309
StdDev	184.564

Class: WX5 (Nom)

Picture6: Visualization of Attributes Weather

ACTORS TABLE:

The Actor table represents the total strength of the actor along with serial number of attacker and serial number of defenders for the particular battle. Most of the time a belligerent is the same as an actor, but sometimes a belligerent consists of several actors. The strength attribute from the actor table attacker strength and defender _strength through Isqno is represented in below chart. The x-axis represents the actor. The Y-axis represents the total strength of the attacker and defender.

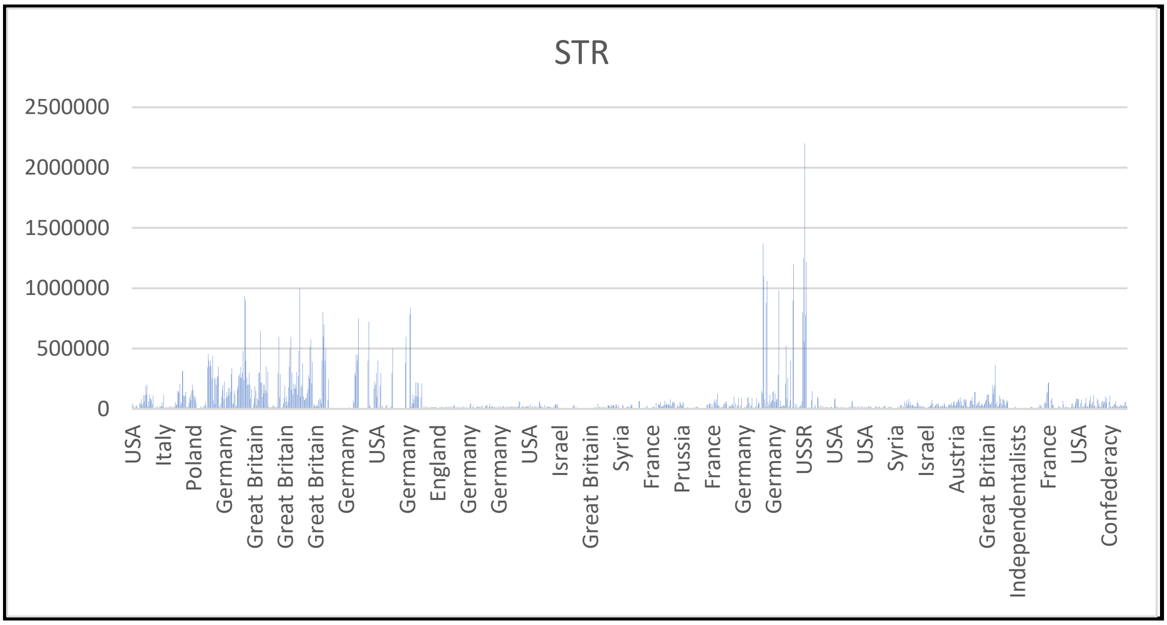


Chart4: STR

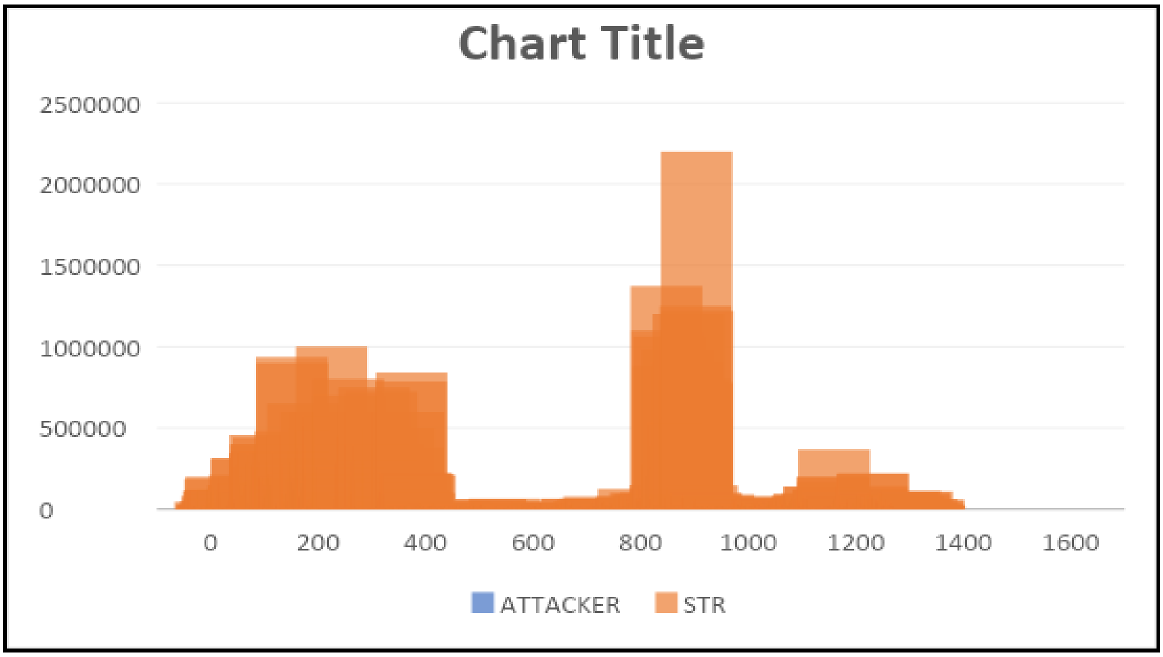
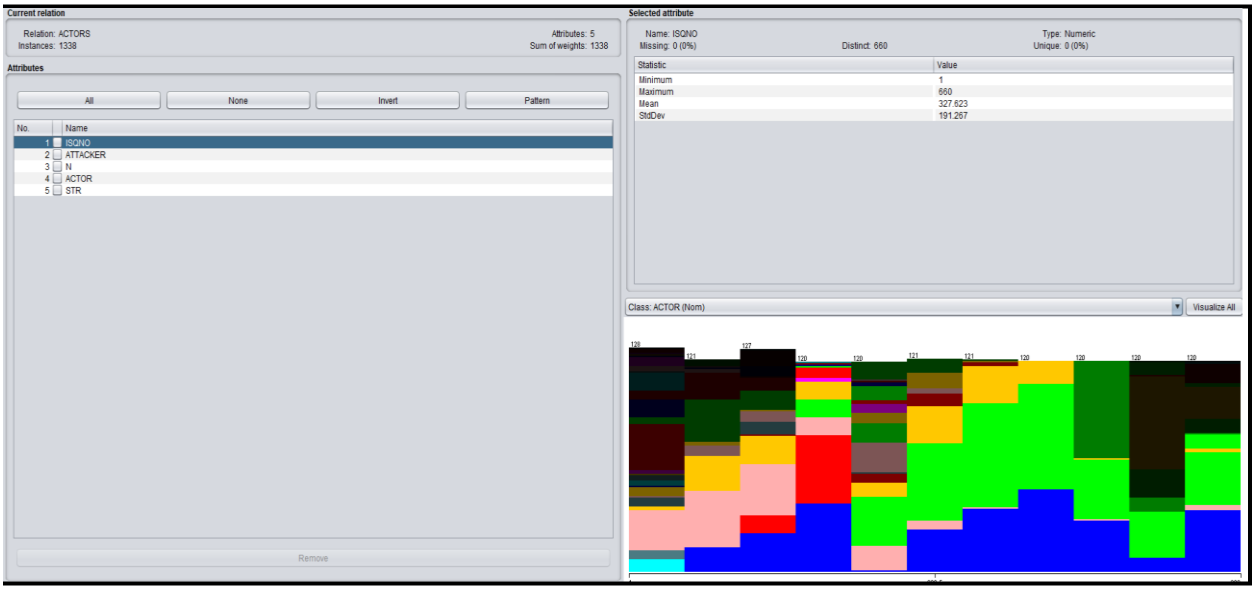


Chart5: Attacker, STR

The SQL query used to create the visualization is :

```
CREATE VIEW ACTORS_TEST AS
SELECT* FROM ACTORS;
```



Current relation

Relation: ACTORS
Instances: 1338

Attributes: 5
Sum of weights: 1338

Attributes

All | None | Invert | Pattern

No.	Name
1	□ ISQNO
2	□ ATTACKER
3	□ N
4	□ ACTOR
5	□ STR

Remove

Selected attribute

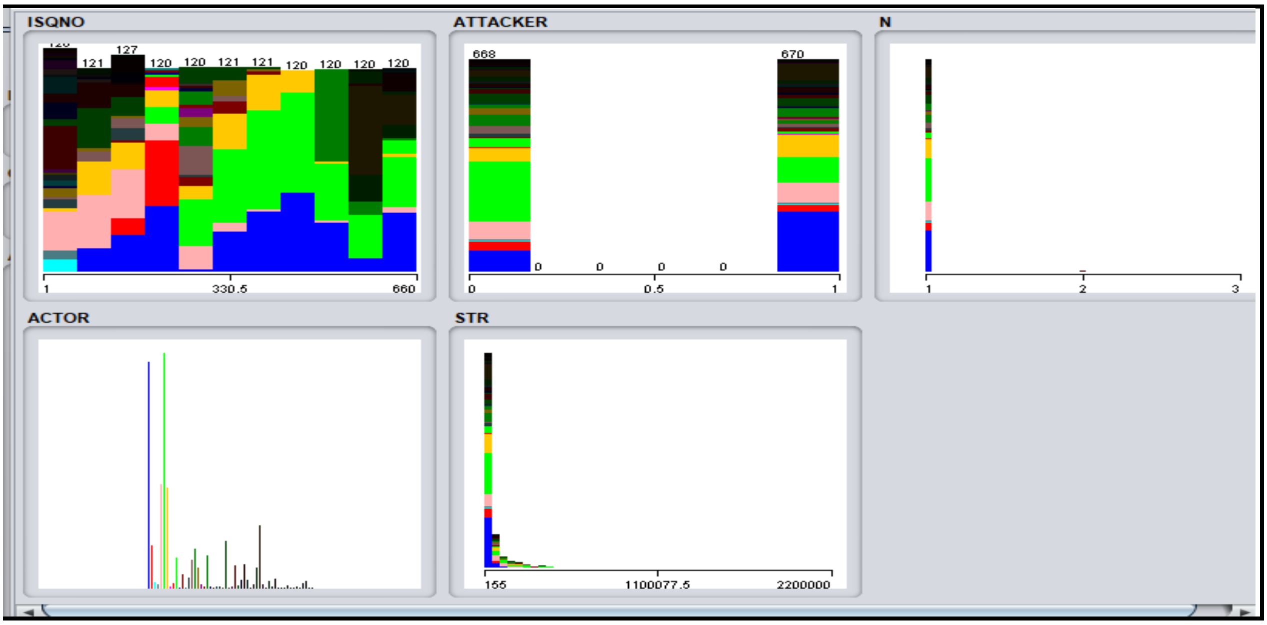
Name: ISQNO
Missing: 0 (0%)
Distinct: 660
Type: Numeric
Unique: 0 (0%)

Statistic	Value
Minimum	1
Maximum	660
Mean	327.623
StdDev	191.267

Class: ACTOR (Nom)

Visualize All

Picture7: Attributes in Actors



Picture8:Visualization on attributes of Actors

FRONT_WIDTHS:

It displays attacker and defender’s front width in kilometres.

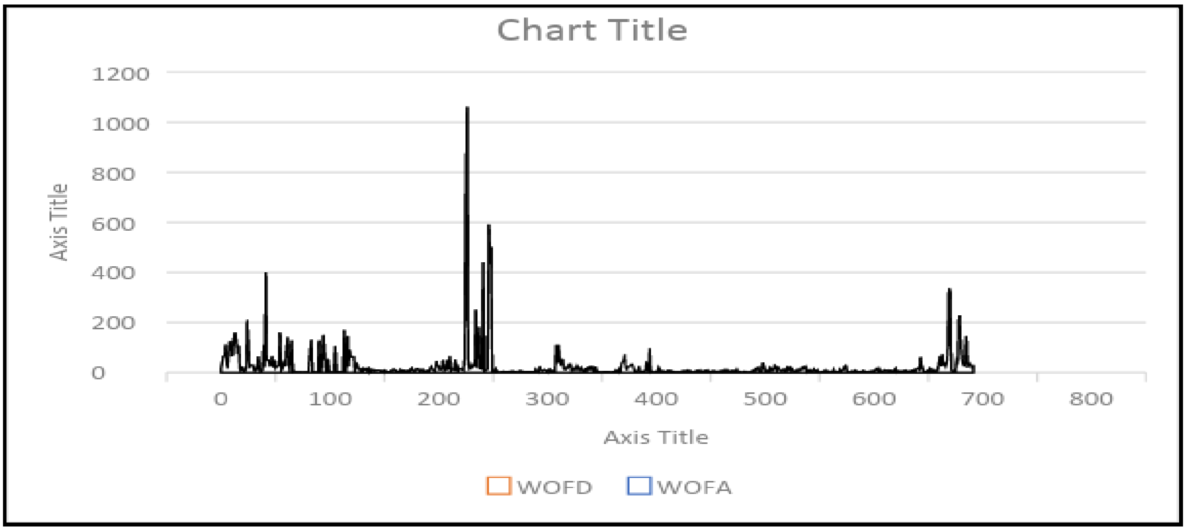
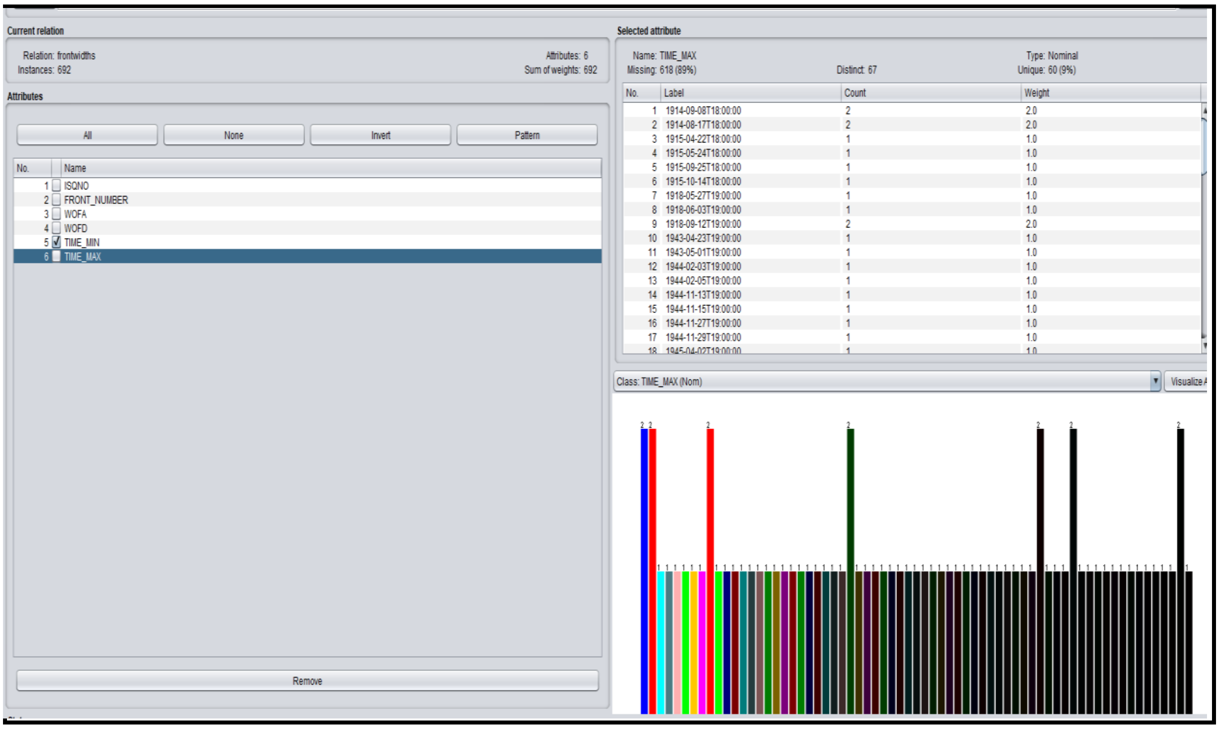


Chart6: WOFD,WOFA

The SQL query used to create the visualization is :

```
CREATE VIEW FRONT_WIDTHS_TEST AS
SELECT* FROM FRONT_WIDTHS ;
```



Relation: frontwidths
Instances: 692
Attributes: 6
Sum of weights: 692

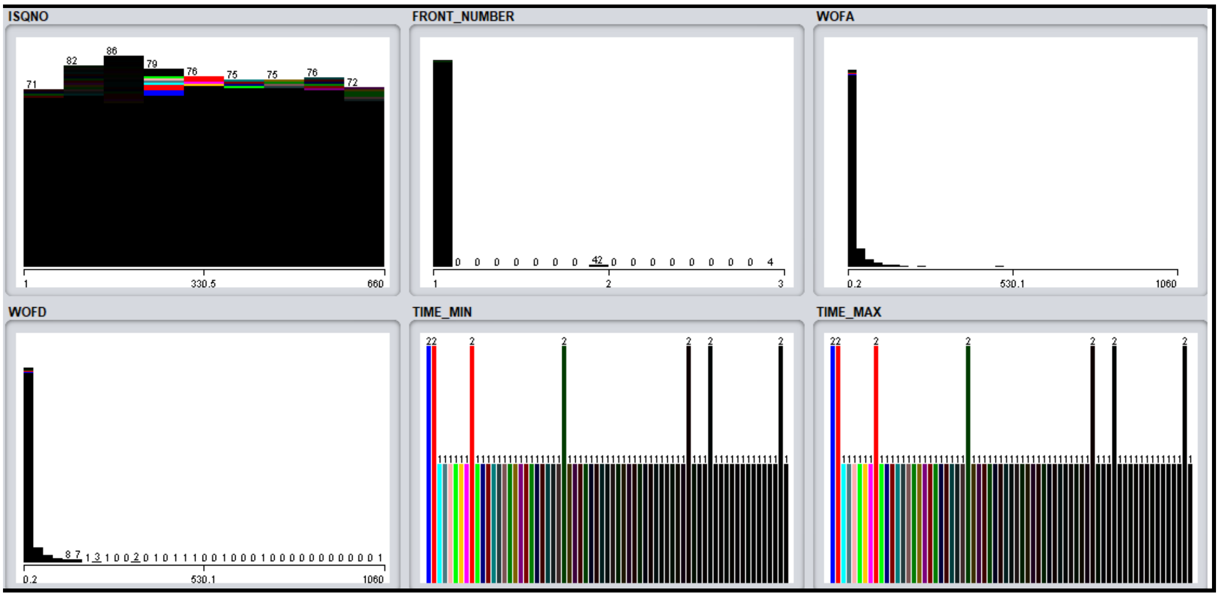
No.		Name
1	□	ISQNO
2	□	FRONT_NUMBER
3	□	WOFA
4	□	WOFD
5	☑	TIME_MIN
6	□	TIME_MAX

Name: TIME_MAX
Missing: 618 (89%)
Distinct: 67
Type: Nominal
Unique: 60 (9%)

No.	Label	Count	Weight
1	1914-09-08T18:00:00	2	2.0
2	1914-08-17T18:00:00	2	2.0
3	1915-04-22T18:00:00	1	1.0
4	1915-05-24T18:00:00	1	1.0
5	1915-09-25T18:00:00	1	1.0
6	1915-10-14T18:00:00	1	1.0
7	1918-05-27T19:00:00	1	1.0
8	1918-06-03T19:00:00	1	1.0
9	1918-09-12T19:00:00	2	2.0
10	1943-04-23T19:00:00	1	1.0
11	1943-05-01T19:00:00	1	1.0
12	1944-02-03T19:00:00	1	1.0
13	1944-02-05T19:00:00	1	1.0
14	1944-11-13T19:00:00	1	1.0
15	1944-11-15T19:00:00	1	1.0
16	1944-11-27T19:00:00	1	1.0
17	1944-11-29T19:00:00	1	1.0
18	1945-04-02T19:00:00	1	1.0

Class: TIME_MAX (Nom)

Picture9: Attributes of Front_Widths



Picture10:Visualization of attributes Front_Widths

TERRAIN:

It describes surface of the land and kind of environment at the time of battle.

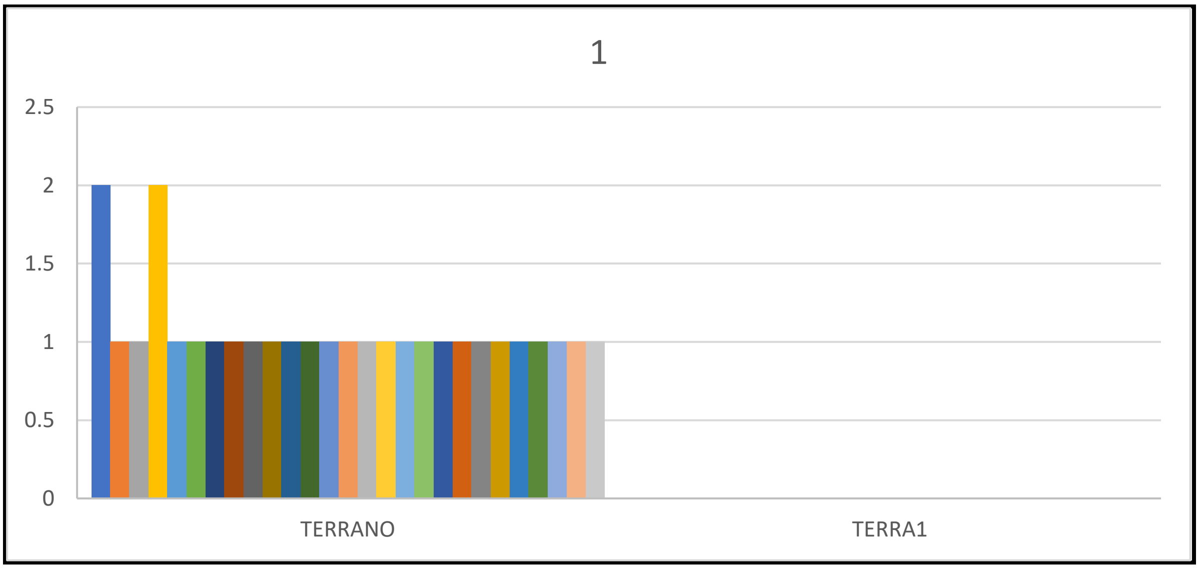
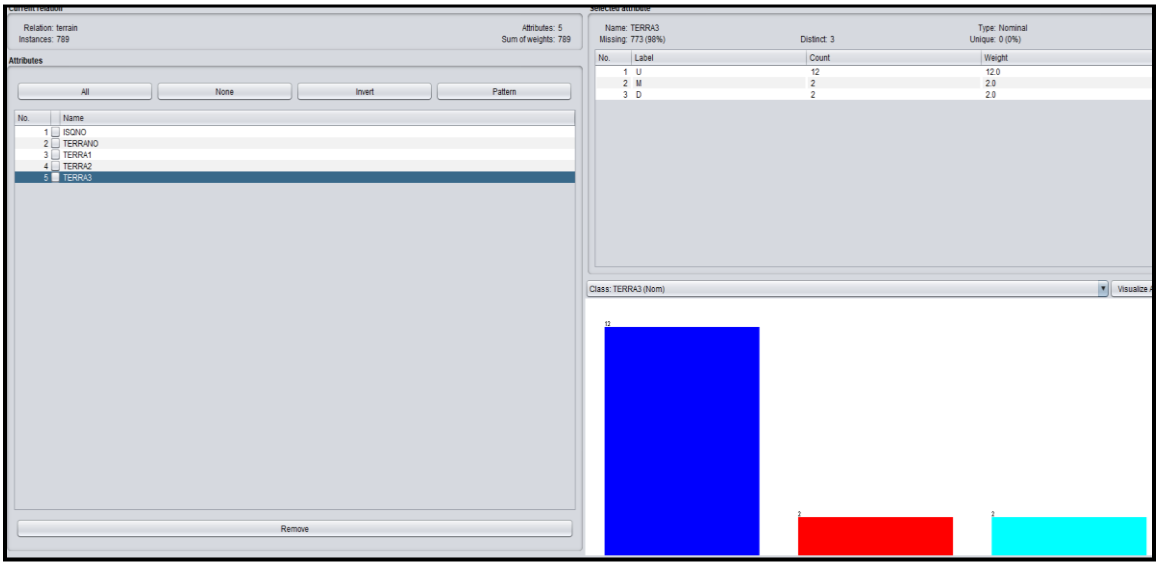


Chart7: Terrano,Terra1

The SQL query used to create the visualization is :

```
CREATE VIEW TERRAIN_TEST AS
SELECT* FROM TERRAIN;
```



Relation: terrain
Instances: 789
Attributes: 5
Sum of weights: 789

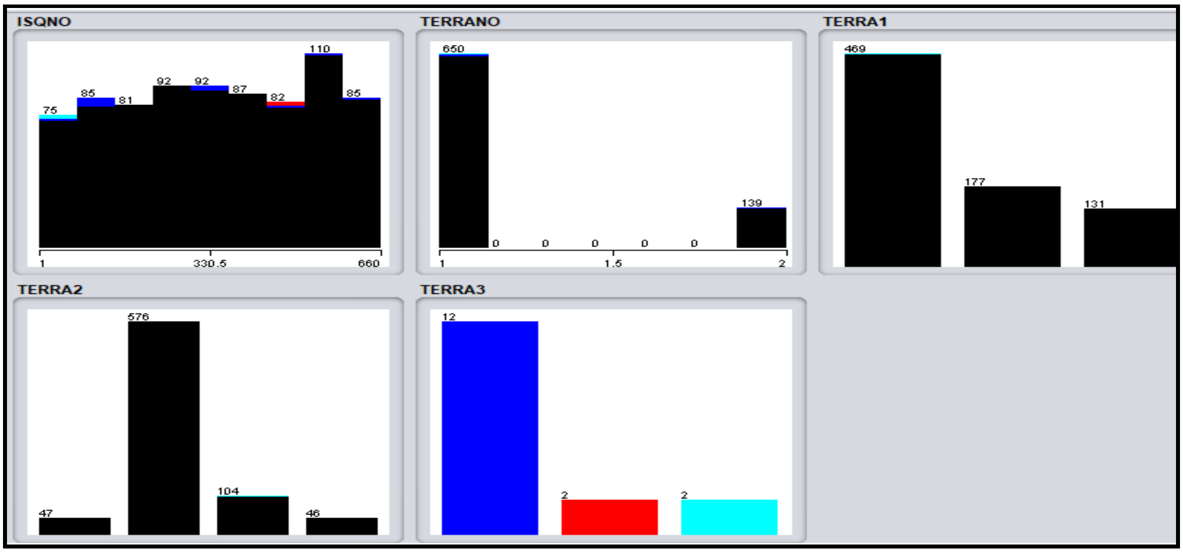
No.		Name
1	□	ISQNO
2	□	TERRANO
3	□	TERRA1
4	□	TERRA2
5	□	TERRA3

Name: TERRA3
Missing: 773 (98%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	U	12	12.0
2	M	2	2.0
3	D	2	2.0

Class: TERRA3 (Nom)

Picture11: Attributes of Terrain



Picture12: Visualization of attributes Terrain

OUTCOME:

The relative achievement of the attacker is predicted through measuring via the difference of attackers and defenders achievement values, as well as to understand which factors contribute to determine the achievement. Achievement is indicated with ach variable and need to predict is Predict ach_attacker-ach_defender.

We need to split the data into two separate views:

```
CREATE OR REPLACE VIEW TRAIN AS
SELECT * FROM TRAIN_TEST
WHERE rel_ach IS NOT NULL;

CREATE OR REPLACE VIEW _TEST
SELECT * FROM TRAIN_TEST
WHERE rel_ach IS NULL;
```

ACHIEVEMENT CHART:

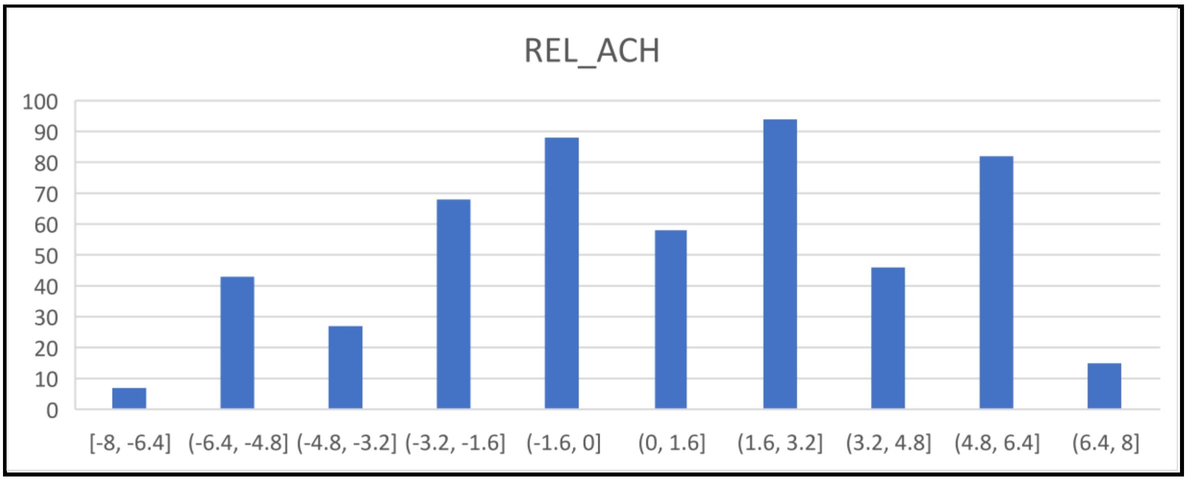


Chart8: Real Achievement is not null denotes x-axis and y-axis

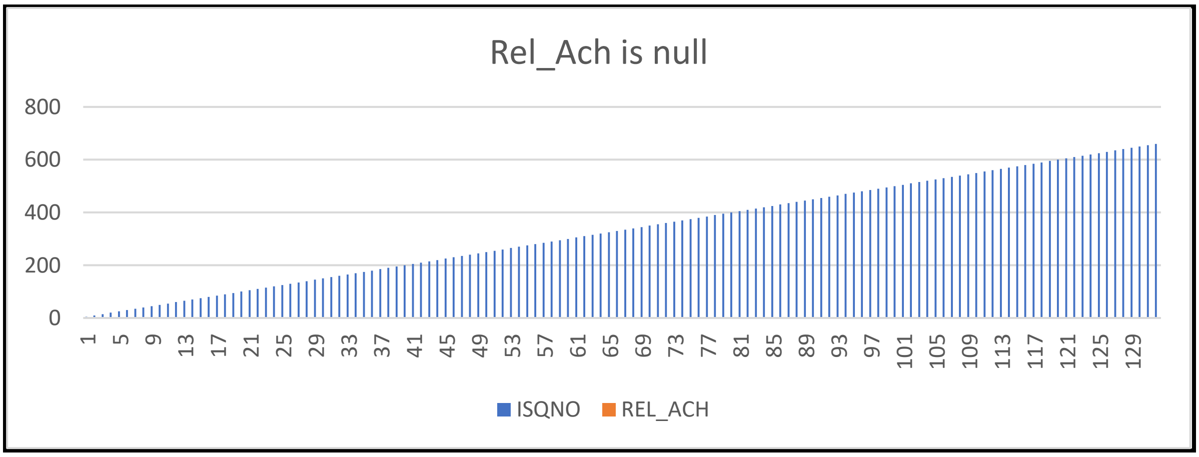


Chart8: Real Achievement is null denotes x-axis and y-axis

PREDICTIVE MODELLING:

Our dataset represents regression data, The predictive modelling target is numeric.

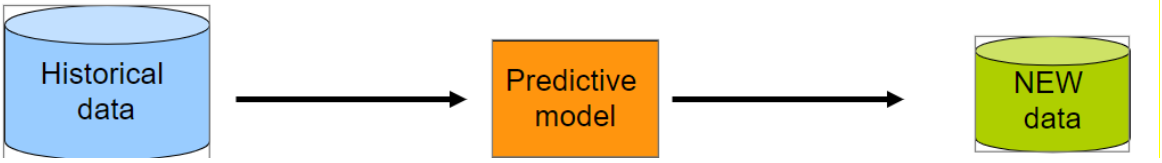
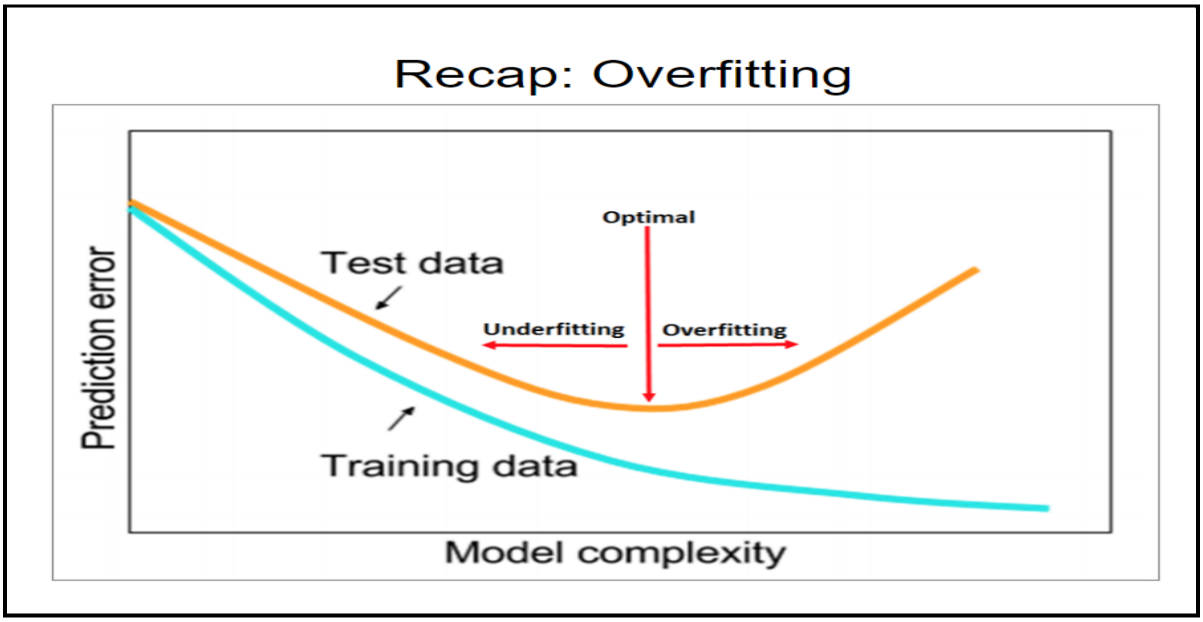


Fig 2: Predictive Modelling

A good model should predict reasonably well on historical data. The main goal of predictive models should predict well on unseen data.



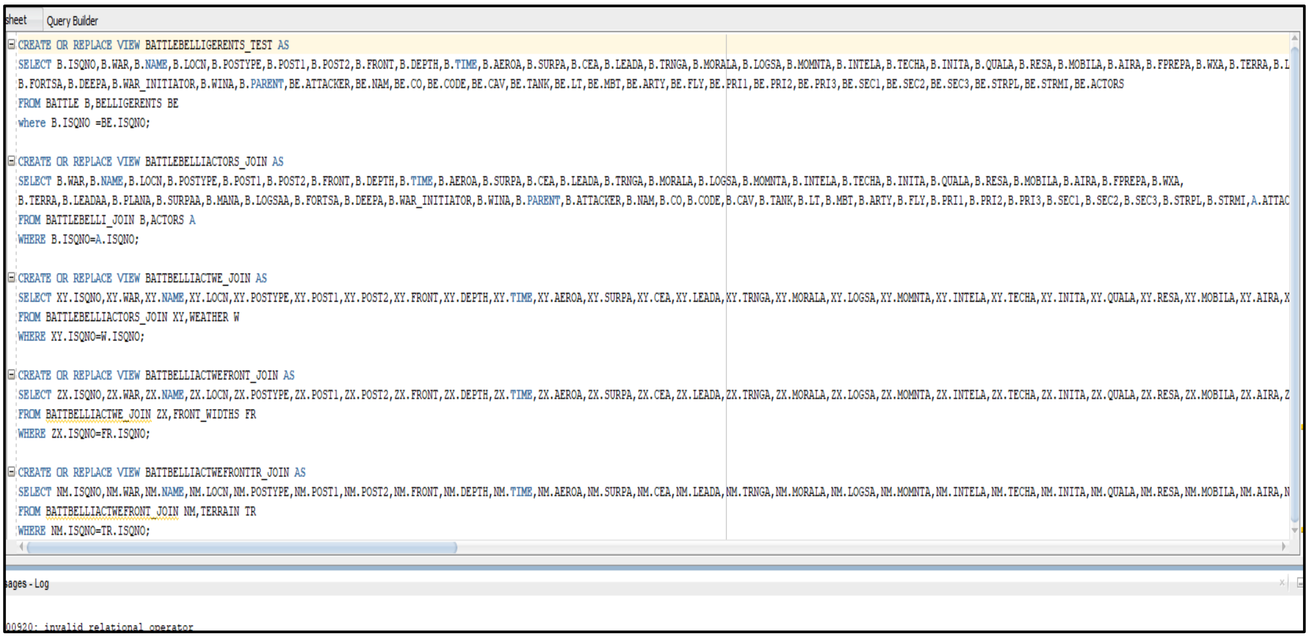
Picture13: Regression modelling

DATA EXPLORATION: Before starting the pre-processing of the data, data exploration can hint what pre-processing is need to be applied and which algorithms can be useful for predictive modelling. The focus is to reduce the error as the model is based on Regression Data. several datasets designed for creating a best predictive model.

DATA MODELLING:

DATASET1:

All the tables are joined together in SQL by creating views. When we joined all 6-tables there are   instances and   attributes. We have created several datasets for creating a best predictive model.



```
CREATE OR REPLACE VIEW BATTLEBELLIGERENTS_TEST AS
SELECT B.ISQNO,B.WAR,B.NAME,B.LOCN,B.POSTYPE,B.POST1,B.POST2,B.FRONT,B.DEPTH,B.TIME,B.AEROA,B.SURPA,B.CEA,B.LEADA,B.TRNGA,B.MORALA,B.LOGSA,B.MOMNTA,B.INTELA,B.TECHA,B.INITA,B.QUALA,B.RESA,B.MOBILA,B.AIRA,B.FPREPA,B.WXA,B.TERRA,B.L
B.FORTSA,B.DEEPA,B.WAR_INITIATOR,B.WINA,B.PARENT,BE.ATTACKER,BE.NAM,BE.CO,BE.CODE,BE.CAV,BE.TANK,BE.LT,BE.MBT,BE.ARTY,BE.FLY,BE.PRI1,BE.PRI2,BE.PRI3,BE.SEC1,BE.SEC2,BE.SEC3,BE.STRPL,BE.STRMI,BE.ACTORS
FROM BATTLE B,BELLIGERENTS BE
where B.ISQNO =BE.ISQNO;

CREATE OR REPLACE VIEW BATTLEBELLIACTORS_JOIN AS
SELECT B.WAR,B.NAME,B.LOCN,B.POSTYPE,B.POST1,B.POST2,B.FRONT,B.DEPTH,B.TIME,B.AEROA,B.SURPA,B.CEA,B.LEADA,B.TRNGA,B.MORALA,B.LOGSA,B.MOMNTA,B.INTELA,B.TECHA,B.INITA,B.QUALA,B.RESA,B.MOBILA,B.AIRA,B.FPREPA,B.WXA,
B.TERRA,B.LEADAA,B.PLANA,B.SURPAA,B.MANA,B.LOGSAA,B.FORTSA,B.DEEPA,B.WAR_INITIATOR,B.WINA,B.PARENT,B.ATTACKER,B.NAM,B.CO,B.CODE,B.CAV,B.TANK,B.LT,B.MBT,B.ARTY,B.FLY,B.PRI1,B.PRI2,B.PRI3,B.SEC1,B.SEC2,B.SEC3,B.STRPL,B.STRMI,A.ATTAC
FROM BATTLEBELLI_JOIN B,ACTORS A
WHERE B.ISQNO=A.ISQNO;

CREATE OR REPLACE VIEW BATTBELLIACTWE_JOIN AS
SELECT XY.ISQNO,XY.WAR,XY.NAME,XY.LOCN,XY.POSTYPE,XY.POST1,XY.POST2,XY.FRONT,XY.DEPTH,XY.TIME,XY.AEROA,XY.SURPA,XY.CEA,XY.LEADA,XY.TRNGA,XY.MORALA,XY.LOGSA,XY.MOMNTA,XY.INTELA,XY.TECHA,XY.INITA,XY.QUALA,XY.RESA,XY.MOBILA,XY.AIRA,X
FROM BATTLEBELLIACTORS_JOIN XY,WEATHER W
WHERE XY.ISQNO=W.ISQNO;

CREATE OR REPLACE VIEW BATTBELLIACTWEFRONT_JOIN AS
SELECT ZX.ISQNO,ZX.WAR,ZX.NAME,ZX.LOCN,ZX.POSTYPE,ZX.POST1,ZX.POST2,ZX.FRONT,ZX.DEPTH,ZX.TIME,ZX.AEROA,ZX.SURPA,ZX.CEA,ZX.LEADA,ZX.TRNGA,ZX.MORALA,ZX.LOGSA,ZX.MOMNTA,ZX.INTELA,ZX.TECHA,ZX.INITA,ZX.QUALA,ZX.RESA,ZX.MOBILA,ZX.AIRA,Z
FROM BATTBELLIACTWE_JOIN ZX,FRONT_WIDTHS FR
WHERE ZX.ISQNO=FR.ISQNO;

CREATE OR REPLACE VIEW BATTBELLIACTWEFRONTTR_JOIN AS
SELECT NM.ISQNO,NM.WAR,NM.NAME,NM.LOCN,NM.POSTYPE,NM.POST1,NM.POST2,NM.FRONT,NM.DEPTH,NM.TIME,NM.AEROA,NM.SURPA,NM.CEA,NM.LEADA,NM.TRNGA,NM.MORALA,NM.LOGSA,NM.MOMNTA,NM.INTELA,NM.TECHA,NM.INITA,NM.QUALA,NM.RESA,NM.MOBILA,NM.AIRA,N
FROM BATTBELLIACTWEFRONT_JOIN NM,TERRAIN TR
WHERE NM.ISQNO=TR.ISQNO;
```

```
00920: invalid relational operator
```

After exporting the CSV file to Weka we have 2069 instances and 75 attributes. Some attributes as these are not helpful for Battle prediction. Then Both training data and test data saved in arff files and prediction file is generated through various algorithms. The regression models are in Weka:----->Classifiers------→Functions

Algorithm	Statistics	Cross-validation	Percentage split	Kaggle score
Linear Regression	Correlation coefficient 0.7401 Mean absolute error 1.9043 Root mean squared error 2.4096 Relative absolute error 64.0347 % Root relative squared error 67.2511 % Total Number of Instances 1679	10	66	4.18045
SMOreg	Correlation coefficient 0.7177 Mean absolute error 1.8212 Root mean squared error 2.5102 Relative absolute error 61.242 % Root relative squared error 70.0596 % Total Number of Instances 1679	10	66	4.97698
Multilayer Perceptron	Correlation coefficient 0.9645 Mean absolute error 0.4332 Root mean squared error 0.9627 Relative absolute error 14.5662 % Root relative squared error 26.8698 % Total Number of Instances 1679	10	66	4.32623
Lazy-iBK	Correlation coefficient 0.7814 Mean absolute error 0.7165 Root mean squared error 2.4037 Relative absolute error 24.0936 % Root relative squared error 67.085 % Total Number of Instances 1679	10	66	4.40627

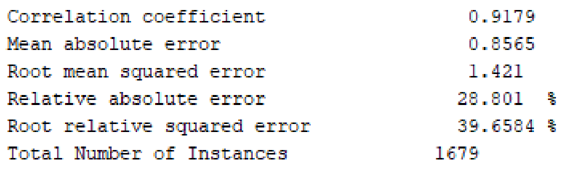
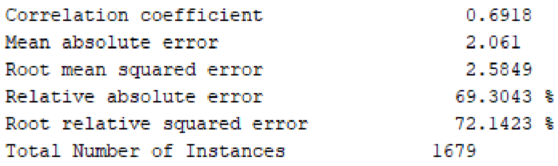
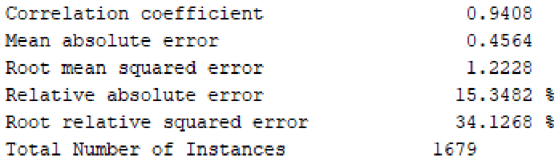
Trees--- **□M5P**	 Correlation coefficient 0.9179 Mean absolute error 0.8565 Root mean squared error 1.421 Relative absolute error 28.801 % Root relative squared error 39.6584 % Total Number of Instances 1679	10	66	5.28467
Trees---- **□Decision Stump**	 Correlation coefficient 0.6918 Mean absolute error 2.061 Root mean squared error 2.5849 Relative absolute error 69.3043 % Root relative squared error 72.1423 % Total Number of Instances 1679	10	66	NA
TREES--- **RANDOM TREE**	 Correlation coefficient 0.9408 Mean absolute error 0.4564 Root mean squared error 1.2228 Relative absolute error 15.3482 % Root relative squared error 34.1268 % Total Number of Instances 1679	10	66	NA

Table 2: Weka Results

FINDINGS:

The results are tried with different cross-validation and percentage split but no change in results occurred. The Weka results were high in Root mean squared error. As the accuracy of predictive model is not well due to over instances and duplicate isqno, Further Data Cleaning Process is followed to prepare a good Training Data for prediction.

DATASET2:

DATA MODELLING:

As Battle is the main table and Belligerent is having important information about Battle sides we combined both of them to get a better predictive model. The SQL query used to generate the Train data and Test data is:

```
CREATE OR REPLACE VIEW NEWBELLIG AS
SELECT ATT.ISQNO, ATT.NAME,ATT.CO AS ATTACKER_COMMANDER,ATT.POSTYPE AS
ATTACKER_POSTYPE,ATT.POST1 AS ATTACKER_POST1,ATT.FRONT AS
ATTACKER_FRONT,ATT.DEPTH AS ATTACKER_DEPTH,ATT.TIME AS
ATTACKER_TIME,ATT.AEROA AS ATTACKER_AEROA,ATT.CEA AS
ATTACKER_CEA,ATT.LEADA AS ATT_LEADA,ATT.TRNGA AS
ATTACKER_TRNGA,ATT.MORALA AS ATTACKER_MORALA,ATT.LOGSA AS
ATTACKER_LOGSA,ATT.MOMNTA AS ATTACKER_MOMNTA,ATT.INTELA AS
ATTACKER_INTELA,ATT.TECHA AS ATTACKER_TECHA,ATT.INITA AS
ATTACKER_INITA,ATT.QUALA AS ATTACKER_QUALA,ATT.RESA AS
ATTACKER_RESA,ATT.MOBILA AS ATTACKER_MOBILA,ATT.AIRA AS
ATTACKER_AIRA,ATT.FPREPA AS ATTACKER_FPREPA,ATT.WXA AS
ATTACKER_WXA,ATT.TERRA AS ATTACKER_TERRA,ATT.LEADAA AS
ATTACKER_LEADAA,ATT.PLANA AS ATTACKER_PLANA,ATT.SURPAA AS
ATTACKER_SURPAA,ATT.MANA AS ATTACKER_MANA,ATT.LOGSAA AS
ATTACKER_LOGSAA,ATT.FORTSA AS ATTACKER_FORTSA,ATT.WAR_INITIATOR AS
ATTACKER_WAR_INITIATOR,DEF.CO AS
DEFENDER_COMMANDER,ATT.CODE AS ATTACKER_CODE,DEF.CODE AS
DEFENDER_CODE,ATT.TANK AS ATTACKER_TANK,DEF.TANK AS DEFENDER_TANK,
ATT.MBT AS ATTACKER_MBT,DEF.MBT AS DEFENDER_MBT,ATT.ARTY AS
```

```
ATTACKER_ARTY,DEF.ARTY AS DEFENDER_ARTY,ATT.FLY AS ATTACKER_FLY,DEF.FLY
AS DEFENDER_FLY,ATT.PRI1 AS ATTACKER_PRI1,DEF.PRI1 AS
DEFENDER_PRI1,ATT.PRI2,ATT.PRI3,ATT.SEC1,ATT.SEC2,ATT.SEC3,ATT.STRPL AS
ATTACKER_STRPL,DEF.STRPL AS DEFENDER_STRPL,ATT.STRMI AS
ATTACKER_STRMI,DEF.STRMI AS DEFENDER_STRMI


FROM
(SELECT BATTLE.ISQNO,
NAME,CO,POSTYPE,POST1,FRONT,DEPTH,TIME,AEROA,CEA,LEADA,TRNGA,MORALA,LOGS
A,MOMNTA,INTELA,QUALA,TECHA,INITA,RESA,MOBILA,AIRA,FPREPA,WXA,TERRA,LEAD
AA,PLANA,SURPAA,MANA,LOGSAA,FORTSA,WAR_INITIATOR,CODE,TANK,MBT,ARTY,FLY,
PRI1,PRI2,PRI3,SEC1,SEC2,SEC3,STRPL,STRMI FROM BATTLE JOIN BELLIGERENTS ON
BATTLE.ISQNO = BELLIGERENTS.ISQNO WHERE ATTACKER=1) ATT
JOIN
(SELECT BATTLE.ISQNO, NAME,
CO,POSTYPE,POST1,FRONT,DEPTH,TIME,AEROA,CEA,LEADA,TRNGA,MORALA,LOGSA,MO
MNTA,INTELA,QUALA,TECHA,INITA,RESA,MOBILA,AIRA,FPREPA,WXA,TERRA,LEADAA,PL
ANA,SURPAA,MANA,LOGSAA,FORTSA,WAR_INITIATOR,CODE,TANK,MBT,ARTY,FLY,PRI1,P
RI2,PRI3,SEC1,SEC2,SEC3,STRPL,STRMI FROM BATTLE JOIN BELLIGERENTS ON
BATTLE.ISQNO = BELLIGERENTS.ISQNO WHERE ATTACKER=0) DEF
ON ATT.ISQNO=DEF.ISQNO
ORDER BY ATT.ISQNO;

SELECT* FROM NEWBELLIG;
CREATE OR REPLACE VIEW PRE12 AS
SELECT NEWBELLIG.*, ATT.ACH-DEF.ACH AS REL_ACH FROM
(SELECT ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=1) ATT JOIN (SELECT
ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=0) DEF
ON ATT.ISQNO=DEF.ISQNO

JOIN NEWBELLIG
ON ATT.ISQNO=NEWBELLIG.ISQNO
ORDER BY ATT.ISQNO;

CREATE OR REPLACE VIEW PREDICTIONNEWTRAIN AS
SELECT * FROM PRE12
WHERE REL_ACH IS NOT NULL;

CREATE OR REPLACE VIEW PREDICTIONNEWTEST AS
SELECT * FROM PRE12
WHERE REL_ACH IS NULL;
```

DATA REDUCTION:

DIMENSIONALITY REDUCTION:

Dimensionality reduction aims at reducing the number of attributes, so that data becomes more focused to the target we want to predict, can improve accuracy and reduce computational time. It Leads to simpler and more interpretable predictive models.

DATA CLEANING:

Attributes are removed from the train data for a better predictive modelling. Entities must not be multiplied beyond necessity. There were 64 attributes and 528 instances from which:ISQNO,NAME,FORTSA,FPREPA,SURPAA,QUALA,RESA,ATTACKER_COMMANDER,DEFENDER_COMMANDER,PR2,PR3,SEC1,SEC2,SEC3 are removed from the dataset. Some of The attributes are removed as they have high percentage of missing values above 80% and some with nominal datatypes as these not helping in prediction. The same attributes are also remove from the Test data to make both Train data and Test Data compatible. Both the files are saved in .ARFF format and then files opened in Text Editor and the headers of predictionnewtest.arff is replaced with predictionnewtrain.arff. Now we applied the model that we have created to test data to generate predictions for the instances where the outcome of the match is unknown. We tested our dataset with various algorithms as mentioned below details:

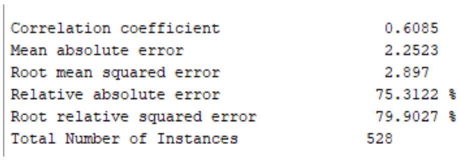
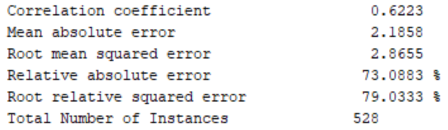
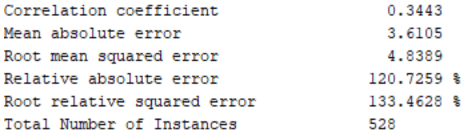
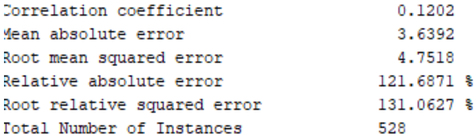
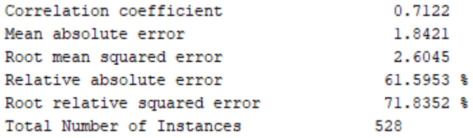
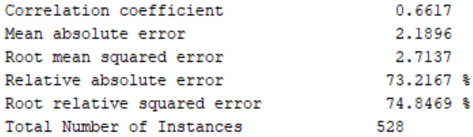
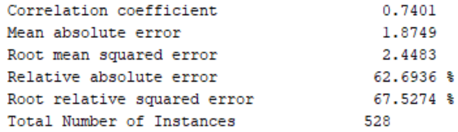
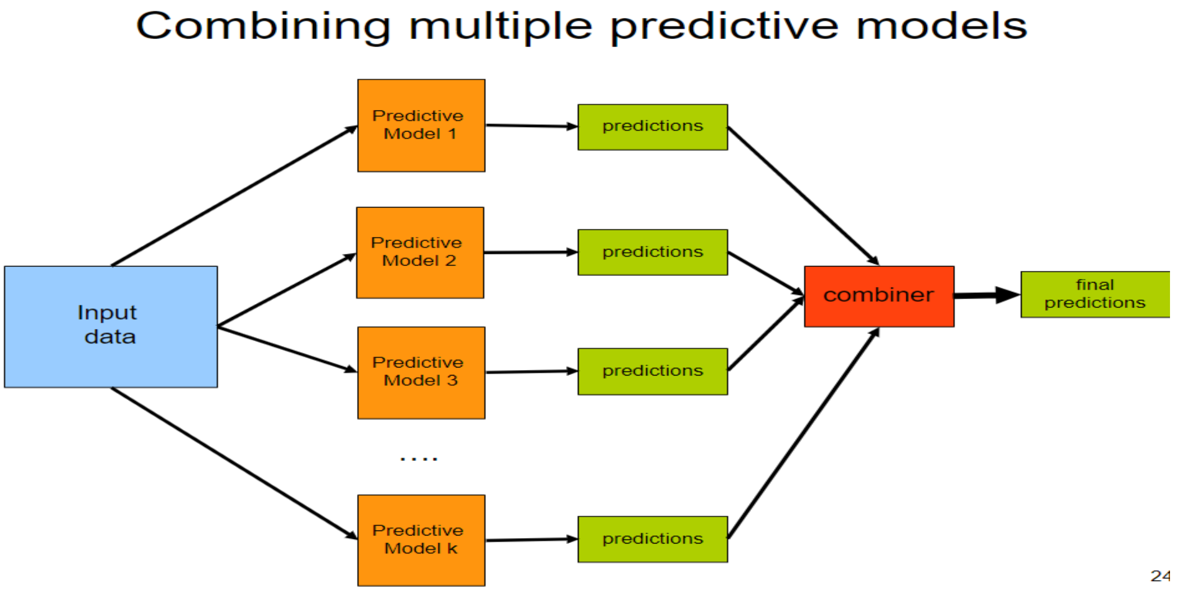
Algorithm	**Statistics**	**Cross-validation**	**Percentage split**	**Kaggle score**
Linear Regression	 Correlation coefficient 0.6085 Mean absolute error 2.2523 Root mean squared error 2.897 Relative absolute error 75.3122 % Root relative squared error 79.9027 % Total Number of Instances 528	10	66	2.27892
SMOreg	 Correlation coefficient 0.6223 Mean absolute error 2.1858 Root mean squared error 2.8655 Relative absolute error 73.0883 % Root relative squared error 79.0333 % Total Number of Instances 528	10	66	2.27892
Multilayer Perceptron	 Correlation coefficient 0.3443 Mean absolute error 3.6105 Root mean squared error 4.8389 Relative absolute error 120.7259 % Root relative squared error 133.4628 % Total Number of Instances 528	10	66	NA
Lazy-iBK	 Correlation coefficient 0.1202 Mean absolute error 3.6392 Root mean squared error 4.7518 Relative absolute error 121.6871 % Root relative squared error 131.0627 % Total Number of Instances 528	10	66	NA
Trees---□M5P	 Correlation coefficient 0.7122 Mean absolute error 1.8421 Root mean squared error 2.6045 Relative absolute error 61.5953 % Root relative squared error 71.8352 % Total Number of Instances 528	10	66	2.16426
Trees----□Decision Stump	 Correlation coefficient 0.6617 Mean absolute error 2.1896 Root mean squared error 2.7137 Relative absolute error 73.2167 % Root relative squared error 74.8469 % Total Number of Instances 528	10	66	2.17260
Rules---□Decision Table	 Correlation coefficient 0.7401 Mean absolute error 1.8749 Root mean squared error 2.4483 Relative absolute error 62.6936 % Root relative squared error 67.5274 % Total Number of Instances 528	10	66	2.19929

Table 3:Weka Results

ENSEMBLE METHOD:

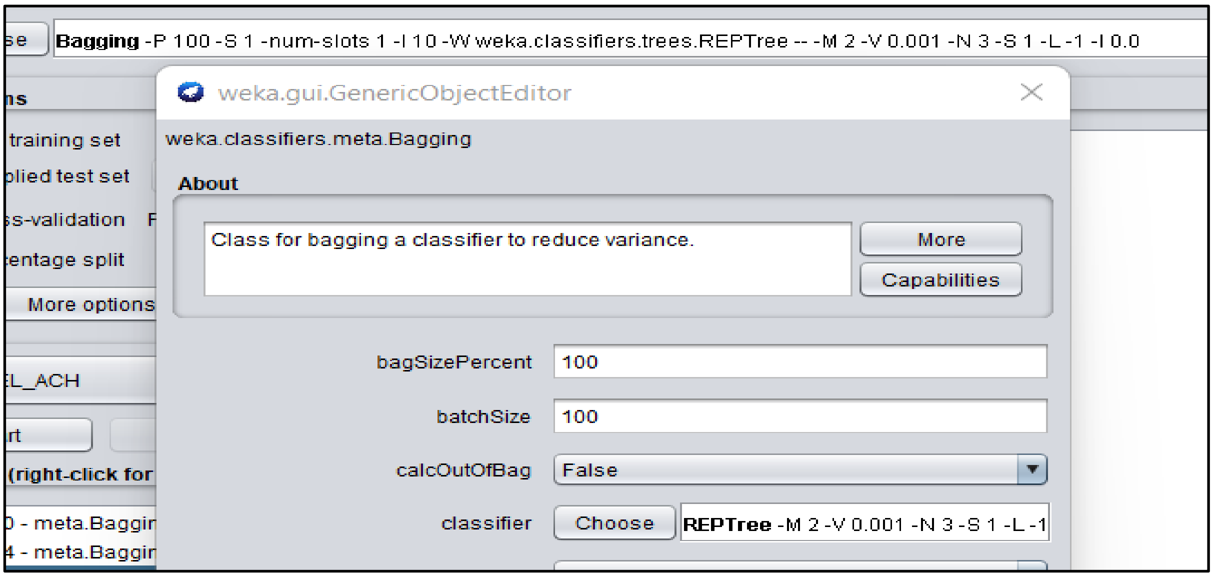
Ensemble methods combine multiple predictive models. It improves predictive performance by Specialization ,Increasing model stability, prevent overfitting.



Picture 14: Multiple Predictive Model

The Ensemble method for Regression Tasks are :

1.Bagging: Bagging is done as (Classifiers -> meta -> Bagging) with regression tree (Classifier -> trees -> REPTree), bagging with a Multilayer Perceptron and linear regression.



Bagging -P 100 -S 1 -num-slots 1 -I 10 -W weka.classifiers.trees.REPTree -- -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0

weka.classifiers.meta.Bagging

Class for bagging a classifier to reduce variance.

bagSizePercent 100

batchSize 100

calcOutOfBag False

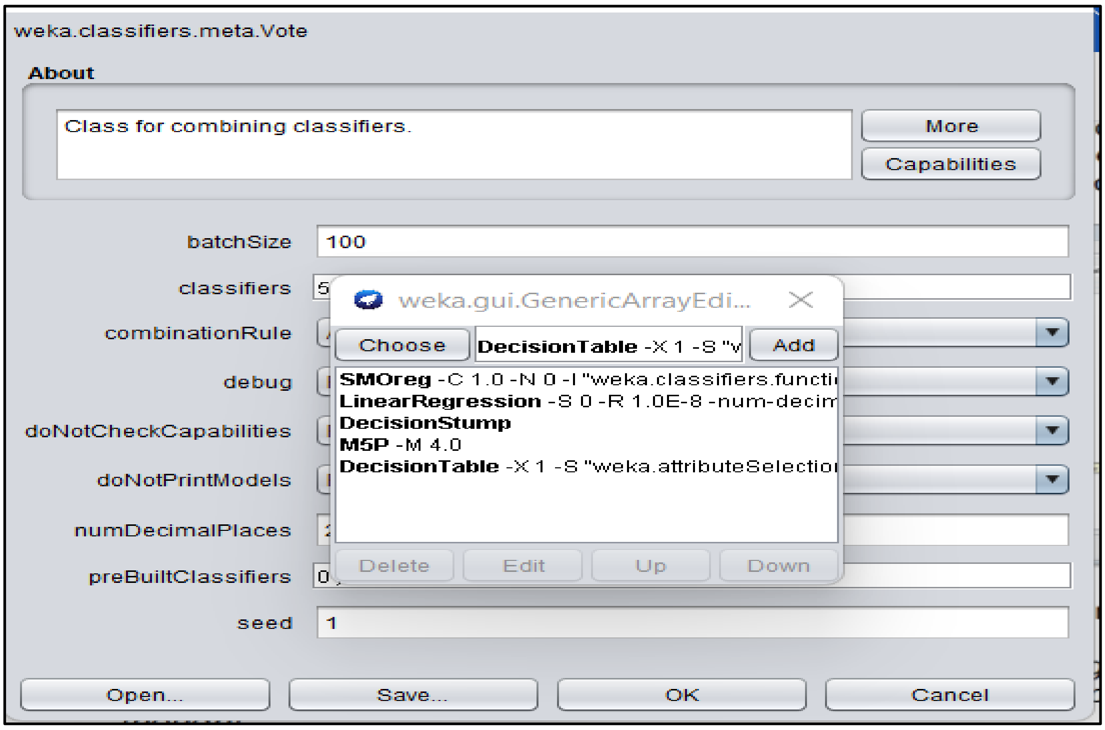
classifier **REPTree** -M 2 -V 0.001 -N 3 -S 1 -L -1

Picture 15: Bagging

2. Random Sub space

3.Voting

Vote Algorithm Process: An ensemble was built combining the five classifiers. Choose: Classifiers -> meta -> Vote



weka.classifiers.meta.Vote

About

Class for combining classifiers.

batchSize: 100

classifiers: 5

seed: 1

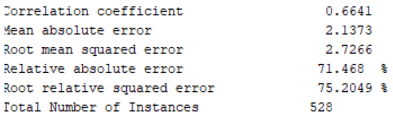
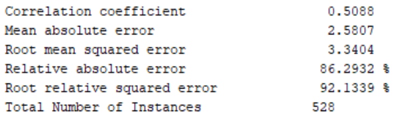
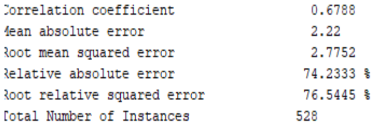
weka.gui.GenericArrayEdi...

DecisionTable -X 1 -S "v

SMOreg -C 1.0 -N 0 -I "weka.classifiers.functi
LinearRegression -S 0 -R 1.0E-8 -num-decim
DecisionStump
M5P -M 4.0
DecisionTable -X 1 -S "weka.attributeSelectio

Picture 16: Voting classifier selection

The algorithm results are mentioned below:

Algorithm	**Statistics**	**Cross Validation**	**Percentage split**	**Kaggle score**
Classifiers -> meta -> Bagging **(Classifier -> trees -> REPTree)**	 Correlation coefficient 0.6641 Mean absolute error 2.1373 Root mean squared error 2.7266 Relative absolute error 71.468 % Root relative squared error 75.2049 % Total Number of Instances 528	10	66	2.38472
Classifiers -> meta -> Bagging **(Classifier -> functions-> Multi Layer Perceptrion)**	 Correlation coefficient 0.5088 Mean absolute error 2.5807 Root mean squared error 3.3404 Relative absolute error 86.2932 % Root relative squared error 92.1339 % Total Number of Instances 528	10	66	NA
Meta-----Random sub space	 Correlation coefficient 0.6788 Mean absolute error 2.22 Root mean squared error 2.7752 Relative absolute error 74.2333 % Root relative squared error 76.5445 % Total Number of Instances 528	10	66	2.43756

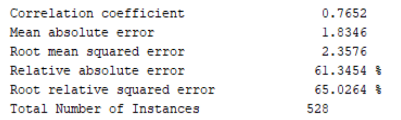
Meta----□Vote	 Correlation coefficient 0.7652 Mean absolute error 1.8346 Root mean squared error 2.3576 Relative absolute error 61.3454 % Root relative squared error 65.0264 % Total Number of Instances 528	10	66	2.10057

Table4: Weka Results

FINDINGS:

The output prediction CSV files are uploaded in Kaggle .Some results are not uploaded as having high Root Mean Squared error .The Prediction makes good score with the Ensemble method -Voting as it worked through 5-different classifiers as SMOrg, Linear Regression, Decision Stump,M5P,Decision Table.

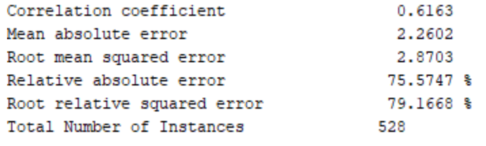
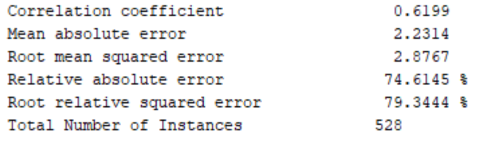
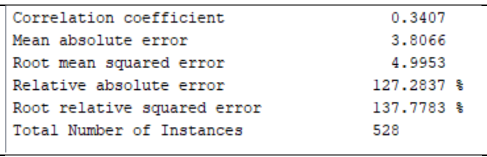
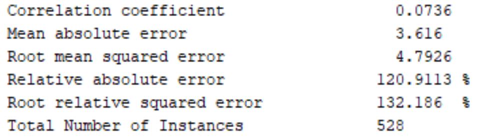
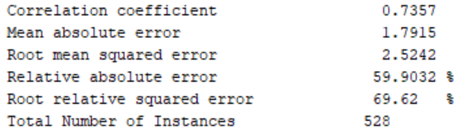
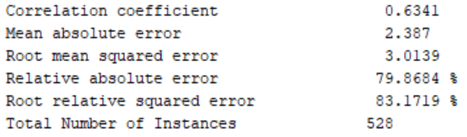
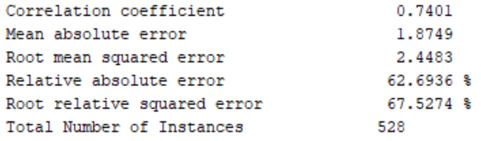
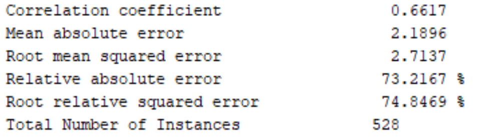
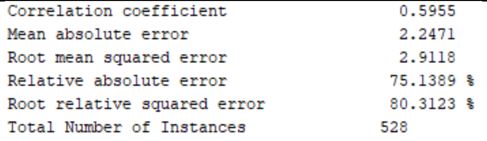
DATASET 3:

The Battle ,Belligerent and weather is joined as as climate ,season ,temperature ,precipitation have a definite impact in war. WXNO =1 is selected to filter the data as this represents the weather condition at the beginning of the battle. The created model was applied to test data to generate predictions for the instances where the outcome of the match is unknown. The dataset were tested with various algorithms as mentioned below details.

```
CREATE OR REPLACE VIEW BATTBELLIWEATHER AS

SELECT NEWBELLIG.*,STT.WX1 AS START_WXA,MID.WX1 AS MIDDLE_WX1,STT.WX2 AS
START_WX2,MID.WX2 AS MIDDLE_WX2,STT.WX3 AS START_WX3, MID.WX3 AS
MIDDLE_WX3,STT.WX4 AS START_WX4, MID.WX4 AS MIDDLE_WX4,STT.WX5 AS
START_WX5, MID.WX5 AS MIDDLE_WX5 FROM

(SELECT ISQNO,WX1,WX2,WX3,WX4,WX5 FROM WEATHER WHERE WXNO=1)STT JOIN
(SELECT ISQNO,WX1,WX2,WX3,WX4,WX5 FROM WEATHER WHERE WXNO=2)MID
ON STT.ISQNO=MID.ISQNO(+)

JOIN NEWBELLIG

ON STT.ISQNO=NEWBELLIG.ISQNO
ORDER BY STT.ISQNO;

CREATE OR REPLACE VIEW BATTBELLW AS
SELECT BBEW.*, ATT.ACH-DEF.ACH AS REL_ACH FROM
(SELECT ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=1) ATT JOIN (SELECT
ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=0) DEF
ON ATT.ISQNO=DEF.ISQNO
JOIN BBEW
ON ATT.ISQNO=BBEW.ISQNO
ORDER BY ATT.ISQNO;
CREATE OR REPLACE VIEW BATTBELLWTRAIN AS
SELECT * FROM  BATTBELLW
WHERE REL_ACH IS NOT NULL;
CREATE OR REPLACE VIEW BATTBELLWTEST AS
SELECT * FROM  BATTBELLW
WHERE REL_ACH IS NULL;
```

DATA CLEANING:

The BATTBELLWTRAIN .CSV has 56 attributes and 528 instances.The attributes ISQNO,FORTSA,DEEPA,WAR_INITIATOR,ATTACKER_COMMANDER,DEFENDER_COMMANDER,CODE,TANK,LT,MBT,ARTY,FLY,PRI2,PRI3,SEC1,SEC2,SEC3,WXNO are removed as they not providing good prediction modelling. The same attributes are also remove from the Test data .The Weka results are represented below:

Algorithm	Statistics	Cross-validation	Percentage split	Kaggle score
Linear Regression	 Correlation coefficient 0.6163 Mean absolute error 2.2602 Root mean squared error 2.8703 Relative absolute error 75.5747 % Root relative squared error 79.1668 % Total Number of Instances 528	10	66	2.23674
SMOreg	 Correlation coefficient 0.6199 Mean absolute error 2.2314 Root mean squared error 2.8767 Relative absolute error 74.6145 % Root relative squared error 79.3444 % Total Number of Instances 528	10	66	NA
Multilayer Perceptron	 Correlation coefficient 0.3407 Mean absolute error 3.8066 Root mean squared error 4.9953 Relative absolute error 127.2837 % Root relative squared error 137.7783 % Total Number of Instances 528	10	66	NA
Lazy-iBK	 Correlation coefficient 0.0736 Mean absolute error 3.616 Root mean squared error 4.7926 Relative absolute error 120.9113 % Root relative squared error 132.186 % Total Number of Instances 528	10	66	NA
Trees-→M5P	 Correlation coefficient 0.7357 Mean absolute error 1.7915 Root mean squared error 2.5242 Relative absolute error 59.9032 % Root relative squared error 69.62 % Total Number of Instances 528	10	66	NA
Trees--→RandomForest	 Correlation coefficient 0.6341 Mean absolute error 2.387 Root mean squared error 3.0139 Relative absolute error 79.8684 % Root relative squared error 83.1719 % Total Number of Instances 528	10	66	NA
Rules-→Decision Table	 Correlation coefficient 0.7401 Mean absolute error 1.8749 Root mean squared error 2.4483 Relative absolute error 62.6936 % Root relative squared error 67.5274 % Total Number of Instances 528	10	66	NA
Trees--→Decision Stump	 Correlation coefficient 0.6617 Mean absolute error 2.1896 Root mean squared error 2.7137 Relative absolute error 73.2167 % Root relative squared error 74.8469 % Total Number of Instances 528	10	66	NA
Trees--→REPTree	 Correlation coefficient 0.5955 Mean absolute error 2.2471 Root mean squared error 2.9118 Relative absolute error 75.1389 % Root relative squared error 80.3123 % Total Number of Instances 528	10	66	2.43756

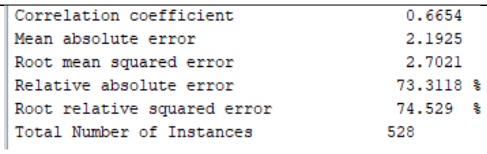
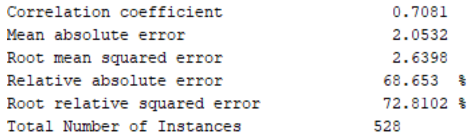
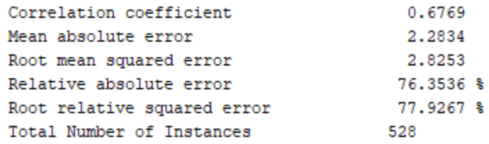
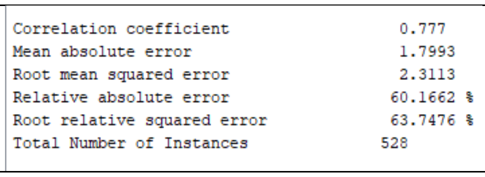
Classifiers -> meta -> Bagging **(Classifier -> trees -> Decision Stump)**	 Correlation coefficient 0.6654 Mean absolute error 2.1925 Root mean squared error 2.7021 Relative absolute error 73.3118 % Root relative squared error 74.529 % Total Number of Instances 528	10	66	2.41890
Classifiers -> meta -> Bagging **(Classifier -> meta-> Decision Table)**	 Correlation coefficient 0.7081 Mean absolute error 2.0532 Root mean squared error 2.6398 Relative absolute error 68.653 % Root relative squared error 72.8102 % Total Number of Instances 528	10	66	2.42430
Meta---→Random sub space	 Correlation coefficient 0.6769 Mean absolute error 2.2834 Root mean squared error 2.8253 Relative absolute error 76.3536 % Root relative squared error 77.9267 % Total Number of Instances 528	10	66	NA
Meta--→Vote	 Correlation coefficient 0.777 Mean absolute error 1.7993 Root mean squared error 2.3113 Relative absolute error 60.1662 % Root relative squared error 63.7476 % Total Number of Instances 528	10	66	2.18272

Table 5: Weka Results Battle+Belligerents+Weather

FINDINGS:

The model provide good result in Rules----→Decision Table and Tree…>Decision stump, but Multilayer Perception and Lazy IBK is giving very high results. Ensemble method Bagging and Vote is done for the data set and Meta……….>Vote give good results as well Kaggle Score. The algorithms chosen for voting are: Decision Tree, Decision Stump,M5P,SMoRG and Linear Regression.

DATASET4:

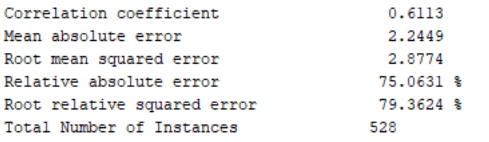
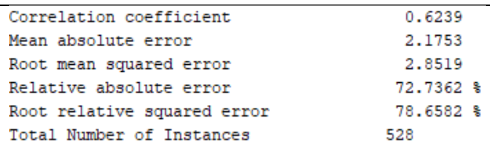
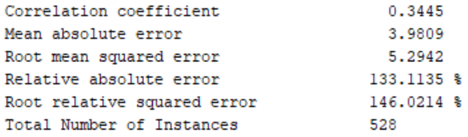
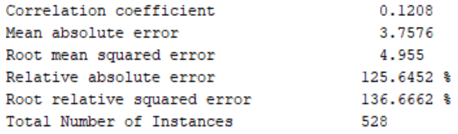
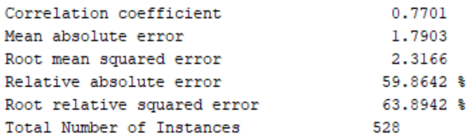
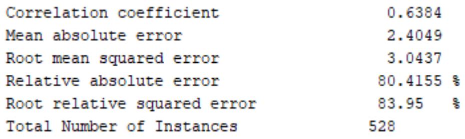
The Battle and Actors is joined as as actors strength have a definite impact in war. The attacker_strength and defender_strength is joined with the main Battle table. The SQL query is mentioned below:

```
CREATE OR REPLACE VIEW BATTLEACTORS AS
SELECT BATTLE.*,ATT.STR AS ATTACKER_STR,DEF.STR AS DEFENDER_STR FROM
(SELECT ISQNO,STR FROM ACTORS WHERE ATTACKER=1)ATT JOIN (SELECT ISQNO,STR
FROM ACTORS WHERE ATTACKER=0)DEF
ON ATT.ISQNO=DEF.ISQNO
JOIN BATTLE
ON ATT.ISQNO=BATTLE.ISQNO
ORDER BY ATT.ISQNO;

CREATE OR REPLACE VIEW  PRED10 AS
SELECT BATTLEACTORS_JOIN .*, ATT.ACH-DEF.ACH AS REL_ACH FROM
(SELECT ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=1) ATT JOIN (SELECT
ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=0) DEF
ON ATT.ISQNO=DEF.ISQNO
```

```
JOIN BATTLEACTORS_JOIN
ON ATT.ISQNO=BATTLEACTORS_JOIN.ISQNO
ORDER BY ATT.ISQNO;
CREATE OR REPLACE VIEW PRED10TRAIN AS
SELECT * FROM  PRED10
WHERE REL_ACH IS NOT NULL;
CREATE OR REPLACE VIEW PRED10TEST AS
SELECT * FROM  PRED10
WHERE REL_ACH IS NULL;
```

DATA CLEANING:

The CSV file has 36 attributes and 528 instances. **ISQNO,POST2,FORTSA** are removed from the Train data. The same attributes are also remove from the Test data. The created model was applied to test data to generate predictions for the instances where the outcome of the match is unknown. The dataset were tested with various algorithms as mentioned below details.

Algorithm	**Statistics**	**Cross-validation**	**Percent age split**	Kaggle score
Linear Regression	 Correlation coefficient 0.6113 Mean absolute error 2.2449 Root mean squared error 2.8774 Relative absolute error 75.0631 % Root relative squared error 79.3624 % Total Number of Instances 528	10	66	NA
SMOreg	 Correlation coefficient 0.6239 Mean absolute error 2.1753 Root mean squared error 2.8519 Relative absolute error 72.7362 % Root relative squared error 78.6582 % Total Number of Instances 528	10	66	NA
Multilayer Perceptron	 Correlation coefficient 0.3445 Mean absolute error 3.9809 Root mean squared error 5.2942 Relative absolute error 133.1135 % Root relative squared error 146.0214 % Total Number of Instances 528	10	66	NA
Lazy-iBK	 Correlation coefficient 0.1208 Mean absolute error 3.7576 Root mean squared error 4.955 Relative absolute error 125.6452 % Root relative squared error 136.6662 % Total Number of Instances 528	10	66	NA
Trees-→M5P	 Correlation coefficient 0.7701 Mean absolute error 1.7903 Root mean squared error 2.3166 Relative absolute error 59.8642 % Root relative squared error 63.8942 % Total Number of Instances 528	10	66	NA
Trees--→RandomForest	 Correlation coefficient 0.6384 Mean absolute error 2.4049 Root mean squared error 3.0437 Relative absolute error 80.4155 % Root relative squared error 83.95 % Total Number of Instances 528	10	66	NA

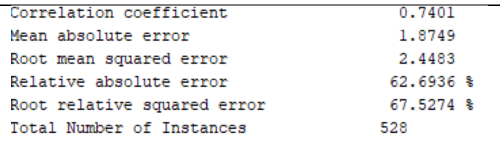
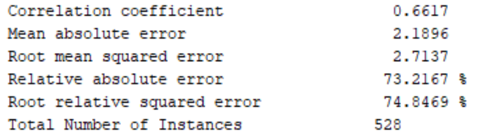
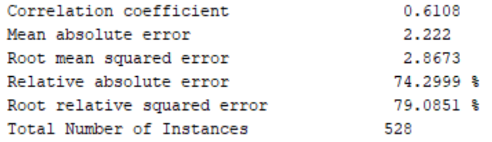
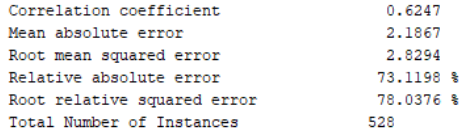
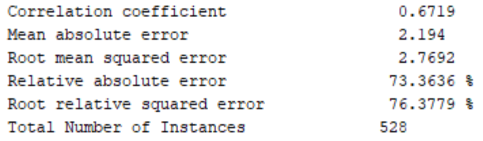
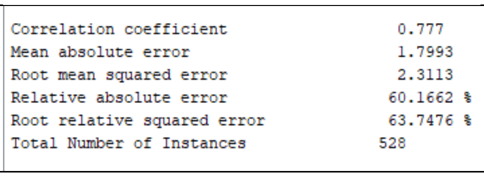
Rules-→Decision Table	Correlation coefficient 0.7401 Mean absolute error 1.8749 Root mean squared error 2.4483 Relative absolute error 62.6936 % Root relative squared error 67.5274 % Total Number of Instances 528	10	66	2.11202
Trees--→Decision Stump	Correlation coefficient 0.6617 Mean absolute error 2.1896 Root mean squared error 2.7137 Relative absolute error 73.2167 % Root relative squared error 74.8469 % Total Number of Instances 528	10	66	NA
Trees--→REPTree	Correlation coefficient 0.6108 Mean absolute error 2.222 Root mean squared error 2.8673 Relative absolute error 74.2999 % Root relative squared error 79.0851 % Total Number of Instances 528	10	66	2.43756
GaussianProcesses	Correlation coefficient 0.6247 Mean absolute error 2.1867 Root mean squared error 2.8294 Relative absolute error 73.1198 % Root relative squared error 78.0376 % Total Number of Instances 528	10	66	2.16426
Meta---→Random sub space	Correlation coefficient 0.6719 Mean absolute error 2.194 Root mean squared error 2.7692 Relative absolute error 73.3636 % Root relative squared error 76.3779 % Total Number of Instances 528	10	66	NA
Meta--→Vote	Correlation coefficient 0.777 Mean absolute error 1.7993 Root mean squared error 2.3113 Relative absolute error 60.1662 % Root relative squared error 63.7476 % Total Number of Instances 528	10	66	2.18272

Table 6: Algorithm for Battle+Actor

FINDINGS:

The model provide good result in Rules---□Decision Table and Tree…>Decision stump, but Multilayer Perception and Lazy IBK is giving very high results. Ensemble method Bagging and Vote is done for the data set and Meta……….>Vote give good results as well Kaggle Score. The algorithms chosen for voting are: Decision Tree, Decision Stump,M5P,SMoRG and Linear Regression.

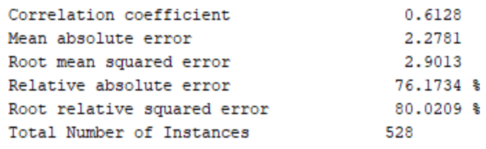
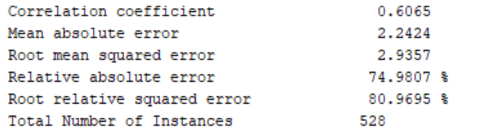
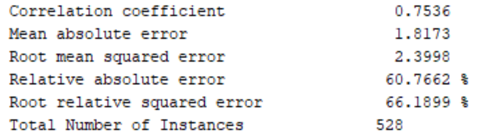
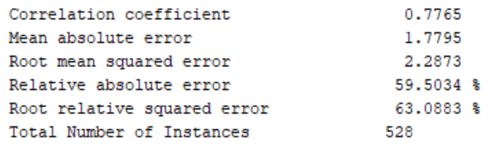
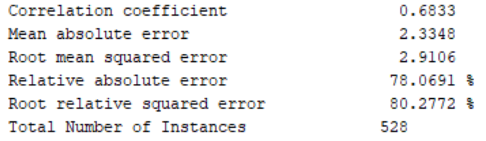
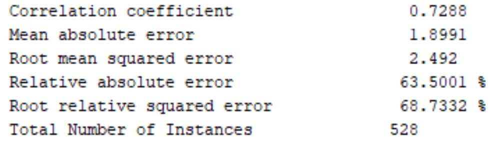
DATASET5:

The Battle,Belligerents and Actors strength are joined together to prepare a good predicative model. The SQL for this is:

```
CREATE OR REPLACE VIEW BATBELACTORS AS
SELECT NEWBELLIG.*,ATT.ATTACKER_STRENGTH AS
ATTACKER_STR,DEF.DEFENDER_STRENGTH AS DEFENDER_STR FROM
(SELECT ISQNO,ATTACKER_STRENGTH FROM STRENGTH_ACTORS )ATT JOIN (SELECT
ISQNO,DEFENDER_STRENGTH FROM STRENGTH_ACTORS)DEF
ON ATT.ISQNO=DEF.ISQNO
JOIN NEWBELLIG
ON ATT.ISQNO=NEWBELLIG.ISQNO
ORDER BY ATT.ISQNO;
```

```
CREATE OR REPLACE VIEW  PRED10 AS
SELECT BATBELACTORS.*, ATT.ACH-DEF.ACH AS REL_ACH FROM
(SELECT DISTINCT ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=1) ATT JOIN
(SELECT DISTINCT ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=0) DEF
ON ATT.ISQNO=DEF.ISQNO
JOIN BATBELACTORS
ON ATT.ISQNO=BATBELACTORS.ISQNO
ORDER BY ATT.ISQNO;
CREATE OR REPLACE VIEW PRED13TRAIN AS
SELECT * FROM  PRED10
WHERE REL_ACH IS NOT NULL;
CREATE OR REPLACE VIEW PRED13TEST AS
SELECT * FROM  PRED10
WHERE REL_ACH IS NULL;
```

The train CSV file has 57 attributes and 528 instances.THE TRAIN CSV FILE was uploaded to Weka and saved in arff format after removing the attributes ISQNO,NAME,FORTSA,FPREPA,SURPAA,QUALA,RESA,ATTACKER_COMMANDER,DEFENDER_COMMANDER,PR2,PR3,SEC1,SEC2,SEC3 and subsequent changed also made to the Test file. The Weka results are mentioned below:

Algorithm	Statistics	Cross-validation	Percentage split	Kaggle score
Linear Regression	 Correlation coefficient 0.6128 Mean absolute error 2.2781 Root mean squared error 2.9013 Relative absolute error 76.1734 % Root relative squared error 80.0209 % Total Number of Instances 528	10	66	NA
SMOreg	 Correlation coefficient 0.6065 Mean absolute error 2.2424 Root mean squared error 2.9357 Relative absolute error 74.9807 % Root relative squared error 80.9695 % Total Number of Instances 528	10	66	NA
Trees-→M5P	 Correlation coefficient 0.7536 Mean absolute error 1.8173 Root mean squared error 2.3998 Relative absolute error 60.7662 % Root relative squared error 66.1899 % Total Number of Instances 528	10	66	NA
Rules-→M5Rules	 Correlation coefficient 0.7765 Mean absolute error 1.7795 Root mean squared error 2.2873 Relative absolute error 59.5034 % Root relative squared error 63.0883 % Total Number of Instances 528	10	66	2.10057
Trees--→Random Forest	 Correlation coefficient 0.6833 Mean absolute error 2.3348 Root mean squared error 2.9106 Relative absolute error 78.0691 % Root relative squared error 80.2772 % Total Number of Instances 528	10	66	NA
Rules-→Decision Table	 Correlation coefficient 0.7288 Mean absolute error 1.8991 Root mean squared error 2.492 Relative absolute error 63.5001 % Root relative squared error 68.7332 % Total Number of Instances 528	10	66	NA

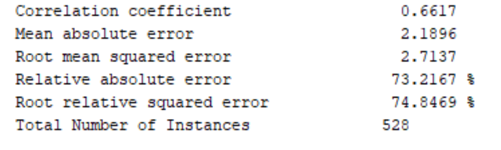
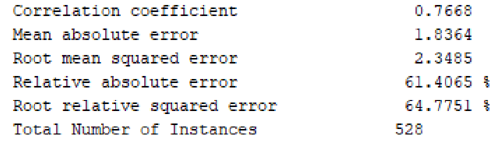
Trees--→Decision Stump	 Correlation coefficient 0.6617 Mean absolute error 2.1896 Root mean squared error 2.7137 Relative absolute error 73.2167 % Root relative squared error 74.8469 % Total Number of Instances 528	10	66	2.11516
Meta--→Vote	 Correlation coefficient 0.7668 Mean absolute error 1.8364 Root mean squared error 2.3485 Relative absolute error 61.4065 % Root relative squared error 64.7751 % Total Number of Instances 528	10	66	2.19192

Table 7: Algorithm results for Battle+Belligerents+Actors

The Kaggle score came as 2.19192 from the above Dataset prediction. So some further Data Reduction and Data Scrubbing Procedure was followed to improve the Kaggle Score.

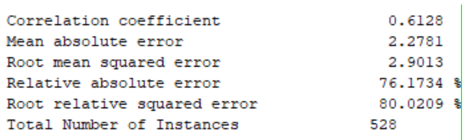
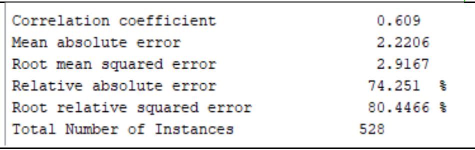
STEP1: To install the relevant package FilteredAttributeEval, we did select Tools->Package Manager from WEKA GUI Chooser. After the Package manager window opens, find the package filteredAttributeSelection and clicked "Install". Confirm the message dialogs.

STEP2: Then reopen WEKA Explorer, reload PRED13TRAIN.arff, go back to Go to the tab "Select attributes" and from Attribute Evaluator menu choose attributeSelection -> FilteredAttribute Eval. WEKA automatically suggested an appropriate search method, accepted the suggestion. Clicked the white space next to where it says Ranker to set the options of the attribute search.

STEP3: When the evaluation finished, we got a list of suggested attributes on the screen .As we are working with a small dataset, you can go back to the Pre-process tab and remove the attributes that were not selected.

STEP4: IMPUTTING MISSING VALUES: In the Pre-process tab choose filters -> unsupervised -> attribute -> ReplaceMissingValues. Click Apply. The dataset is saved in arff and opened it in Notepad. We checked that there are no more missing values.

STEP5: After removing the same attributes from Test data and doing the manual changes in Notepad the arff files were used for prediction.

Algorithm	Statistics	Cross-validation	Percent age split	Kaggle score
Linear Regression	 Correlation coefficient 0.6128 Mean absolute error 2.2781 Root mean squared error 2.9013 Relative absolute error 76.1734 % Root relative squared error 80.0209 % Total Number of Instances 528	10	66	2.10557
SMOreg	 Correlation coefficient 0.609 Mean absolute error 2.2206 Root mean squared error 2.9167 Relative absolute error 74.251 % Root relative squared error 80.4466 % Total Number of Instances 528	10	66	NA

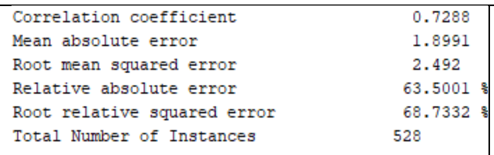
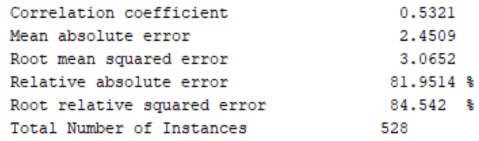
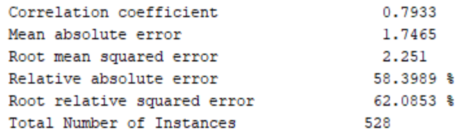
Rules-→Decision Table	Correlation coefficient 0.7288 Mean absolute error 1.8991 Root mean squared error 2.492 Relative absolute error 63.5001 % Root relative squared error 68.7332 % Total Number of Instances 528	10	66	2.17260
Trees--→Decision Stump	Correlation coefficient 0.5321 Mean absolute error 2.4509 Root mean squared error 3.0652 Relative absolute error 81.9514 % Root relative squared error 84.542 % Total Number of Instances 528	10	66	NA
Meta----□Vote	Correlation coefficient 0.7933 Mean absolute error 1.7465 Root mean squared error 2.251 Relative absolute error 58.3989 % Root relative squared error 62.0853 % Total Number of Instances 528	10	66	2.06173

Table 8: Weka Results of Battle+Belligerents+Actors

FINDINGS:

Combining the top performing models into one algorithm using the Bagging and Vote classifiers decreased Relative Absolute error and improved Kaggle Public Leader board Score and DATASET 5 is proved as the best model for the predicative modelling.

EVALUATION AND DEPLOYMENT:

- **EVALUATION:** The objective of the project is to predict relative achievement of the attacker, which is achieved. Rules-→M5Rules, Trees--→Decision Stump, Trees-→M5P Rules-→Decision Table, SMOreg has given Weka results with less Relative Absolute error The final Decision Table configuration proved to be the best for the successful modelling. The Battle table and strength of Belligerents as STRPL and STRMI with Actors Strength has proved as a best model. Overall simple relational table worked better as compared with larger dataset and complex relations. Though the process is a Black model, the business requirement is met and the outcome is predicted with less error . The Linear Regression model also recommended as well as it proved good in simple relational tables.
- **LIMITATIONS:**
  The generated Dataset is one of the limitations. As previously stated, the model was developed with minimal errors. Further analysis of the dataset might be done by adding additional features and accurately replacing missing values to reduce error. The potential for a least error regression model is high, the cost of starting a new project is a major risk the business may face. To run more advanced classifiers, powerful hardware needs to be installed and maintained by technical specialists.
- **FURTHER ACTIONS:**
  There should be actions for adjustment of the models. There should be a focus for proper predictive model either implementing ensemble machine learning algorithms or identifying more potential applicable algorithms.
- **DEPLOYMENT:** The system's deployment is easy and uncomplicated and it does not need any coding skills. The Free software WEKA was installed in the computers and preconfigured models loaded for prediction.

CONCLUSION:

The project is based on the gaming industry and this is a strategy game in which predicting the values of relative achievement of the attacker. Success was measured by model effectiveness. During evaluation, the best model with its specific attributes were concluded. The results are tried with different cross-validation and percentage split and various algorithms Data Underfitting/overfitting was avoided and got good results in the Kaggle Score.

REFERENCES:

1. Chapman, Adam. (2016). Digital Games as History: How Videogames Represent the Past and Offer.
2. Frank, E., Hall, M, A.,and Written, I, H.,2011. Data Mining : Practical Machine Learning Tools and Techniques.3rd Edition.Burlington: Morgan Kaufmann Publishers.
3. Kaggle(2021), Aleksey Bilogur, Historical Military Battles dataset, retrieved from https://www.kaggle.com/residentmario/database-of-battles. (accessed on 19/09/2021).
4. Raessens, Joost, and Jeffrey Goldstein. (2005). Handbook of Computer Game Studies. Cambridge, MA: MIT Press.
5. Recall, Precision,F1-Score: Slides 42-44 (optional).
6. University of Waikato,1999-2016, Weka. Version 3.8.1[computer program]. New Zealand: University of Waikato.
7. WEKA book Chapters 2 and 5 (2nd & 3rd edition).
8. X.Wu,V.Kumar,J.Quinlan,J.ghosh,Q.Yang,H.Motoda,G.McLachlan,A.Ng,B.Liu,Ph.Yu,Z.Zhou, M. Steinbach, D. Hand and D. Steinberg.14(2008). Top 10 algorithms in data mining. Journal Knowledge and Information Systems, 1: 1-37.

########################APPENDIX SQL QUERIES####################

####BATTLE+BELLIGERENTS WITH ALL ATTACKERS AND DEFENDERS OF BELLIGERENT###########

```
CREATE OR REPLACE VIEW NEWBELLIG AS

SELECT ATT.ISQNO, ATT.NAME,ATT.CO AS ATTACKER_COMMANDER,ATT.POSTYPE AS ATTACKER_POSTYPE,ATT.POST1 AS ATTACKER_POST1,ATT.FRONT AS ATTACKER_FRONT,ATT.DEPTH AS ATTACKER_DEPTH,ATT.TIME AS ATTACKER_TIME,ATT.AEROA AS ATTACKER_AEROA,ATT.CEA AS ATTACKER_CEA,ATT.LEADA AS ATT_LEADA,ATT.TRNGA AS ATTACKER_TRNGA,ATT.MORALA AS ATTACKER_MORALA,ATT.LOGSA AS ATTACKER_LOGSA,ATT.MOMNTA AS ATTACKER_MOMNTA,ATT.INTELA AS ATTACKER_INTELA,ATT.TECHA AS ATTACKER_TECHA,ATT.INITA AS ATTACKER_INITA,ATT.QUALA AS ATTACKER_QUALA,ATT.RESA AS ATTACKER_RESA,ATT.MOBILA AS ATTACKER_MOBILA,ATT.AIRA AS ATTACKER_AIRA,ATT.FPREPA AS ATTACKER_FPREPA,ATT.WXA AS ATTACKER_WXA,ATT.TERRA AS ATTACKER_TERRA,ATT.LEADAA AS ATTACKER_LEADAA,ATT.PLANA AS ATTACKER_PLANA,ATT.SURPAA AS ATTACKER_SURPAA,ATT.MANA AS ATTACKER_MANA,ATT.LOGSAA AS ATTACKER_LOGSAA,ATT.FORTSA AS ATTACKER_FORTSA,ATT.WAR_INITIATOR AS ATTACKER_WAR_INITIATOR,DEF.CO AS

DEFENDER_COMMANDER,ATT.CODE AS ATTACKER_CODE,DEF.CODE AS DEFENDER_CODE,ATT.TANK AS ATTACKER_TANK,DEF.TANK AS DEFENDER_TANK, ATT.MBT AS ATTACKER_MBT,DEF.MBT AS DEFENDER_MBT,ATT.ARTY AS ATTACKER_ARTY,DEF.ARTY AS DEFENDER_ARTY,ATT.FLY AS ATTACKER_FLY,DEF.FLY AS DEFENDER_FLY,ATT.PRI1 AS ATTACKER_PRI1,DEF.PRI1 AS DEFENDER_PRI1,ATT.PRI2,ATT.PRI3,ATT.SEC1,ATT.SEC2,ATT.SEC3,ATT.STRPL AS ATTACKER_STRPL,DEF.STRPL AS DEFENDER_STRPL,ATT.STRMI AS ATTACKER_STRMI,DEF.STRMI AS DEFENDER_STRMI



FROM

(SELECT BATTLE.ISQNO, NAME,CO,POSTYPE,POST1,FRONT,DEPTH,TIME,AEROA,CEA,LEADA,TRNGA,MORALA,LOGSA,MOMNTA,INTELA,QUALA,TECHA,INITA,RESA,MOBILA,AIRA,FPREPA,WXA,TERRA,LEADAA,PLANA,SURPAA,MANA,LOGSAA,FORTSA,WAR_INITIATOR,CODE,TANK,MBT,ARTY,FLY,PRI1,PRI2,PRI3,SEC1,SEC2,SEC3,STRPL,STRMI FROM BATTLE JOIN BELLIGERENTS ON BATTLE.ISQNO = BELLIGERENTS.ISQNO WHERE ATTACKER=1) ATT

JOIN

(SELECT BATTLE.ISQNO, NAME, CO,POSTYPE,POST1,FRONT,DEPTH,TIME,AEROA,CEA,LEADA,TRNGA,MORALA,LOGSA,MOMNTA,INTELA,QUALA,TECHA,INITA,RESA,MOBILA,AIRA,FPREPA,WXA,TERRA,LEADAA,PLANA,SURPAA,MANA,LOGSAA,FORTSA,WAR_INITIATOR,CODE,TANK,MBT,ARTY,FLY,PRI1,PRI2,PRI3,SEC1,SEC2,SEC3,STRPL,STRMI FROM BATTLE JOIN BELLIGERENTS ON BATTLE.ISQNO = BELLIGERENTS.ISQNO WHERE ATTACKER=0) DEF

ON ATT.ISQNO=DEF.ISQNO

ORDER BY ATT.ISQNO;
```

```
########################BATTLE+BELLIGERENTS+ACTORS JOIN###########################

Create or replace view bellig as

SELECT att.isqno, att.name,att.postype,att.post1,att.post2,att.front,att.depth,att.time,att.aeroa,att.surpa,att.cea,att.leada,att.trnga,att.morala,att.logsa,att.momnta,att.intela,att.techa,att.inita,att.quala,att.resa,att.mobila,att.aira,att.fprepa,att.wxa,att.terra,att.leadaa,att.plana,att.surpaa,att.mana,att.logsaa,att.fortsa,att.deepa,att.war_initiator,att.co AS attacker_commander, def.co AS

defender_commander,att.code,att.tank,att.lt,att.mbt,att.arty,att.fly,att.pri1,att.pri2,att.pri3,att.sec1,att.sec2,att.sec3,att.strpl,att.strmi FROM




(SELECT battle.isqno,name,postype,post1,post2,front,depth,time,aeroa,surpa,cea,leada,trnga,morala,logsa,momnta,intela,techa,inita,quala,resa,mobila,aira,fprepa,wxa,terra,leadaa,plana,surpaa,mana,logsaa,fortsa,deepa,war_initiator,co,code,tank,lt,mbt,arty,fly,pri1,pri2,pri3,sec1,sec2,sec3,strpl,strmi from battle JOIN belligerents ON battle.isqno =

belligerents.isqno WHERE attacker=1) att

JOIN

(SELECT battle.isqno,name,postype,post1,post2,front,depth,time,aeroa,surpa,cea,leada,trnga,morala,logsa,momnta,intela,techa,inita,quala,resa,mobila,aira,fprepa,wxa,terra,leadaa,plana,surpaa,mana,logsaa,fortsa,deepa,war_initiator,co,code,tank,lt,mbt,arty,fly,pri1,pri2,pri3,sec1,sec2,sec3,strpl,strmi ach from battle JOIN belligerents ON battle.isqno =

belligerents.isqno WHERE attacker=0) def

ON att.isqno=def.isqno

order by att.isqno;


create or replace view strength_actors as

SELECT att.isqno, attacker_strength, defender_strength FROM

(SELECT isqno, SUM(str) AS attacker_strength FROM actors WHERE attacker=1
GROUP BY isqno) att

JOIN (SELECT isqno, SUM(str) AS defender_strength FROM actors WHERE

attacker=0 GROUP BY isqno) def

ON att.isqno=def.isqno

order by isqno;


##################################################################################

CREATE OR REPLACE VIEW BELLIGACTORS_JOIN AS

SELECT
BB.ISQNO,BB.POSTYPE,BB.POST1,BB.POST2,BB.FRONT,BB.DEPTH,BB.TIME,BB.AEROA,BB.SURPA,BB.CEA,BB.LEADA,BB.TRNGA,BB.MORALA,BB.LOGSA,BB.MOMNTA,BB.INTELA,BB.TECHA,BB.INITA,BB.QUALA,BB.RESA,BB.MOBILA,BB.AIRA,BB.FPREPA,BB.WXA,BB.TERRA,BB.LEADAA,BB.PLANA,BB.SURPAA,BB.MANA,BB.LOGSAA,BB.FORTSA,BB.DEEPA,BB.WAR_INITIATOR,BB.ATTACKER_COMMANDER,BB.DEFENDER_COMMANDER,BB.CODE,BB.TANK,BB.LT,BB.MBT,BB.ARTY,BB.FLY,BB.PRI1,BB.PRI2,BB.PRI3,BB.SEC1,BB.SEC2,BB.SEC3,BB.STRPL,BB.STRMI,A.attacker_strength,A.defender_strength

FROM Bellig BB,strength_actors A

WHERE BB.ISQNO=A.ISQNO

order by isqno;



##########################Battle+Belligerent########################################



Create or replace view newbellig as

SELECT att.isqno,
att.name,att.postype,att.post1,,att.front,att.depth,att.time,att.aeroa,att.surpa,att.cea,att.leada,att.trnga,att.morala,att.logsa,att.momnta,att.intela,att.techa,att.inita,att.quala,att.resa,att.mobila,att.aira,att.fprepa,att.wxa,att.terra,att.leadaa,att.plana,att.surpaa,att.mana,att.logsaa,att.fortsa,,att.war_initiator,att.co AS attacker_commander, def.co ASdefender_commander,def.co AS defender_commander,att.code as attacker_code,def.code as defender_code,att.tank as attacker_tank,def.tank as defender_tank, att.mbt as attacker_mbt,def.mbt as defender_mbt,att.arty,att.fly,att.pri1,att.pri2,att.pri3,att.sec1,att.sec2,att.sec3,att.strpl as attacker_strpl,def.strpl as defender_strpl,att.strmi as attacker_strmi,def.strmi as defender_strmi FROM

(SELECT battle.isqno,
name,co,postype,post1,front,depth,time,aeroa,surpa,cea,leada,trnga,morala,logsa,momnta,intela,quala,techa,inita,resa,mobila,aira,fprepa,wxa,terra,leadaa,plana,surpaa,mana,logsaa,fortsa,war_initiator,code,tank,mbt,arty,fly,pri1,pri2,pri3,sec1,sec2,sec3,strpl,strmi from battle JOIN belligerents ON battle.isqno = belligerents.isqno WHERE attacker=1) att

JOIN
```

```
(SELECT battle.isqno, name, co,postype,post1,front,depth,time,aeroa,surpa,cea,leada,trnga,morala,logsa,momnta,intela,quala,techa,inita,resa,mobila,aira,fprepa,wxa,terra,leadaa,plana,surpaa,mana,logsaa,fortsa,war_initiator,code,tank,mbt,arty,fly,pri1,pri2,pri3,sec1,sec2,sec3,strpl,strmi from battle JOIN belligerents ON battle.isqno = belligerents.isqno WHERE attacker=0) def

ON att.isqno=def.isqno

ORDER BY att.isqno;


####################Battle+Belligerent+weather########################

CREATE OR REPLACE VIEW battbelliweather AS


SELECT newbellig.*,stt.wx1 as start_wxa,mid.wx1 as middle_wx1,stt.wx2 as start_wx2,mid.wx2 as middle_wx2,stt.wx3 as start_wx3, mid.wx3 as middle_wx3,stt.wx4 as start_wx4, mid.wx4 as middle_wx4,stt.wx5 as start_wx5, mid.wx5 as middle_wx5 FROM


(SELECT isqno,wx1,wx2,wx3,wx4,wx5 from weather WHERE WXNO=1)stt JOIN (SELECT isqno,wx1,wx2,wx3,wx4,wx5 from weather WHERE WXNO=2)mid


ON stt.isqno=mid.isqno(+)

JOIN newbellig

ON stt.isqno=newbellig.isqno

ORDER BY stt.isqno;


#####################Battle+Belligerent+Actors####################################

Create or replace view newbellig as

SELECT att.isqno, att.name,att.postype,att.post1,,att.front,att.depth,att.time,att.aeroa,att.surpa,att.cea,att.leada,att.trnga,att.morala,att.logsa,att.momnta,att.intela,att.techa,att.inita,att.quala,att.resa,att.mobila,att.aira,att.fprepa,att.wxa,att.terra,att.leadaa,att.plana,att.surpaa,att.mana,att.logsaa,att.fortsa,,att.war_initiator,att.co AS attacker_commander, def.co ASdefender_commander,def.co AS defender_commander,att.code as attacker_code,def.code as defender_code,att.tank as attacker_tank,def.tank as defender_tank, att.mbt as attacker_mbt,def.mbt as defender_mbt,att.arty,att.fly,att.pri1,att.pri2,att.pri3,att.sec1,att.sec2,att.sec3,att.strpl as attacker_strpl,def.strpl as defender_strpl,att.strmi as attacker_strmi,def.strmi as defender_strmi FROM


FROM

(SELECT battle.isqno, name,co,postype,post1,front,depth,time,aeroa,surpa,cea,leada,trnga,morala,logsa,momnta,intela,q
uala,techa,inita,resa,mobila,aira,fprepa,wxa,terra,leadaa,plana,surpaa,mana,logsaa,fortsa,war_initia
tor,code,tank,mbt,arty,fly,pri1,pri2,pri3,sec1,sec2,sec3,strpl,strmi from battle JOIN belligerents ON
battle.isqno =

belligerents.isqno WHERE attacker=1) att

JOIN

(SELECT battle.isqno, name,
co,postype,post1,front,depth,time,aeroa,surpa,cea,leada,trnga,morala,logsa,momnta,intela,quala,te
cha,inita,resa,mobila,aira,fprepa,wxa,terra,leadaa,plana,surpaa,mana,logsaa,fortsa,war_initiator,cod
e,tank,mbt,arty,fly,pri1,pri2,pri3,sec1,sec2,sec3,strpl,strmi from battle JOIN belligerents ON
battle.isqno =

belligerents.isqno WHERE attacker=0) def

ON att.isqno=def.isqno

ORDER BY att.isqno;


CREATE OR REPLACE VIEW battbelliactors AS

SELECT newbellig.*,att.str as attacker_str,def.str as defender_str FROM

(SELECT isqno,str from actors WHERE attacker=1)att JOIN (SELECT isqno,str from actors WHERE
attacker=0)def

ON att.isqno=def.isqno

JOIN newbellig

ON att.isqno=newbellig.isqno

ORDER BY att.isqno;


##############################BATTLE+ACTORS############################


CREATE OR REPLACE VIEW  PREDOMAS AS

SELECT battleactors.*,att.strpl as attacker_strpl,def.strpl as defender_strpl,att.strmi as
attacker_strmi,def.strmi as defender_strmi,ATT.ACH-DEF.ACH AS REL_ACH FROM

(SELECT ISQNO,STRPL,STRMI,ACH FROM BELLIGERENTS WHERE ATTACKER=1) ATT JOIN (SELECT
ISQNO,STRPL,STRMI, ACH FROM BELLIGERENTS WHERE ATTACKER=0) DEF

ON ATT.ISQNO=DEF.ISQNO

JOIN battleactors

ON ATT.ISQNO=battleactors.ISQNO

ORDER BY ATT.ISQNO;

CREATE OR REPLACE VIEW OMTRAIN AS

SELECT * FROM PREDOMAS
```

```
WHERE REL_ACH IS NOT NULL;
CREATE OR REPLACE VIEW OMTEST AS
SELECT * FROM PREDOMAS
WHERE REL_ACH IS NULL;


#################BATTLE+BELLIGERENT+ACTORS+FRONTWIDTH(WOFA,WOFD)###############
CREATE OR REPLACE VIEW BATBELACTORFR   AS
SELECT BATBELACTORS.*,ATT.wofa AS ATTACKER_wofa,DEF.wofd AS DEFENDER_wofd FROM
(SELECT distinct ISQNO,wofa,wofd FROM front_widths WHERE front_number=1)ATT JOIN (SELECT distinct ISQNO,wofa,wofd FROM front_widths WHERE front_number=3)DEF
ON ATT.ISQNO(+)=DEF.ISQNO(+)
JOIN BATBELACTORS
ON ATT.ISQNO=BATBELACTORS.ISQNO
ORDER BY ATT.ISQNO;

#########BATTLE+BELLIGERENT+FRONT_WITHS+TERRAIN+ACTORS#########################
CREATE OR REPLACE VIEW BATBELACTORFR  AS
SELECT
BB.ISQNO,BB.POSTYPE,BB.POST1,BB.POST2,BB.FRONT,BB.DEPTH,BB.TIME,BB.AEROA,BB.SURPA,BB.CEA,BB.LEADA,BB.TRNGA,BB.MORALA,BB.LOGSA,BB.MOMNTA,BB.INTELA,BB.TECHA,BB.INITA,BB.QUALA,BB.RESA,BB.MOBILA,BB.AIRA,BB.FPREPA,BB.WXA,BB.TERRA,BB.LEADAA,BB.PLANA,BB.SURPAA,BB.MANA,BB.LOGSAA,BB.FORTSA,BB.DEEPA,BB.WAR_INITIATOR,BB.WOFA,BB.WOFD,BB.TIME_MIN,BB.TIME_MAX FROM BATTLEFRONT BB,FRONT_WIDTHS F
WHERE BB.ISQNO=F.ISQNO
AND FRONT_NUMBER=1
ORDER BY ISQNO;

CREATE OR REPLACE VIEW  PRED11 AS
SELECT BATBELACTORFR.*, ATT.ACH-DEF.ACH AS REL_ACH FROM
(SELECT DISTINCT ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=1) ATT JOIN (SELECT DISTINCT ISQNO, ACH FROM BELLIGERENTS WHERE ATTACKER=0) DEF
ON ATT.ISQNO=DEF.ISQNO
JOIN BATBELACTORFR
```

```
ON ATT.ISQNO=BATBELACTORFR.ISQNO

ORDER BY ATT.ISQNO;

CREATE OR REPLACE VIEW PRED14TRAIN AS

SELECT * FROM  PRED11

WHERE REL_ACH IS NOT NULL;

CREATE OR REPLACE VIEW PRED14TEST AS

SELECT * FROM  PRED11

WHERE REL_ACH IS NULL;

#######################################################################################

CREATE OR REPLACE VIEW BATTBELLIACTWE_JOIN AS
SELECT
XY.ISQNO,XY.WAR,XY.NAME,XY.LOCN,XY.POSTYPE,XY.POST1,XY.POST2,XY.FRONT,XY.DEPTH,XY.TI
ME,XY.AEROA,XY.SURPA,XY.CEA,XY.LEADA,XY.TRNGA,XY.MORALA,XY.LOGSA,XY.MOMNTA,XY.INT
ELA,XY.TECHA,XY.INITA,XY.QUALA,XY.RESA,XY.MOBILA,XY.AIRA,XY.FPREPA,XY.WXA,XY.TERRA,XY.L
EADAA,XY.PLANA,XY.SURPAA,XY.MANA,XY.LOGSAA,XY.FORTSA,XY.DEEPA,XY.WAR_INITIATOR,XY.
WINA,XY.PARENT,XY.ATTACKER,XY.NAM,XY.CO,XY.CODE,XY.CAV,XY.TANK,XY.LT,XY.MBT,XY.ARTY,
XY.FLY,XY.PRI1,XY.PRI2,XY.PRI3,XY.SEC1,XY.SEC2,XY.SEC3,XY.STRPL,XY.STRMI,XY.ACTORS,XY.N,XY.
ACTOR,XY.STR,W.WXNO,W.WX1,W.WX2,W.WX3,W.WX4,W.WX5
FROM BATTLEBELLIACTORS_JOIN XY,WEATHER W
WHERE XY.ISQNO=W.ISQNO;

CREATE OR REPLACE VIEW BATTBELLIACTWE_JOIN AS
SELECT
XY.ISQNO,XY.WAR,XY.NAME,XY.LOCN,XY.POSTYPE,XY.POST1,XY.POST2,XY.FRONT,XY.DEPTH,XY.TI
ME,XY.AEROA,XY.SURPA,XY.CEA,XY.LEADA,XY.TRNGA,XY.MORALA,XY.LOGSA,XY.MOMNTA,XY.INT
ELA,XY.TECHA,XY.INITA,XY.QUALA,XY.RESA,XY.MOBILA,XY.AIRA,XY.FPREPA,XY.WXA,XY.TERRA,XY.L
EADAA,XY.PLANA,XY.SURPAA,XY.MANA,XY.LOGSAA,XY.FORTSA,XY.DEEPA,XY.WAR_INITIATOR,XY.
WINA,XY.PARENT,XY.ATTACKER,XY.NAM,XY.CO,XY.CODE,XY.CAV,XY.TANK,XY.LT,XY.MBT,XY.ARTY,
XY.FLY,XY.PRI1,XY.PRI2,XY.PRI3,XY.SEC1,XY.SEC2,XY.SEC3,XY.STRPL,XY.STRMI,XY.ACTORS,XY.N,XY.
ACTOR,XY.STR,W.WXNO,W.WX1,W.WX2,W.WX3,W.WX4,W.WX5
FROM BATTLEBELLIACTORS_JOIN XY,WEATHER W
WHERE XY.ISQNO=W.ISQNO;

CREATE OR REPLACE VIEW BATTBELLIACTWEFRONT_JOIN AS
SELECT
ZX.ISQNO,ZX.WAR,ZX.NAME,ZX.LOCN,ZX.POSTYPE,ZX.POST1,ZX.POST2,ZX.FRONT,ZX.DEPTH,ZX.T
IME,ZX.AEROA,ZX.SURPA,ZX.CEA,ZX.LEADA,ZX.TRNGA,ZX.MORALA,ZX.LOGSA,ZX.MOMNTA,ZX.IN
TELA,ZX.TECHA,ZX.INITA,ZX.QUALA,ZX.RESA,ZX.MOBILA,ZX.AIRA,ZX.FPREPA,ZX.WXA,ZX.TERRA,ZX
.LEADAA,ZX.PLANA,ZX.SURPAA,ZX.MANA,ZX.LOGSAA,ZX.FORTSA,ZX.DEEPA,ZX.WAR_INITIATOR,Z
X.WINA,ZX.PARENT,ZX.ATTACKER,ZX.NAM,ZX.CO,ZX.CODE,ZX.CAV,ZX.TANK,ZX.LT,ZX.MBT,ZX.ART
Y,ZX.FLY,ZX.PRI1,ZX.PRI2,ZX.PRI3,ZX.SEC1,ZX.SEC2,ZX.SEC3,ZX.STRPL,ZX.STRMI,ZX.ACTORS,ZX.N,Z
X.ACTOR,ZX.STR,ZX.WXNO,ZX.WX1,ZX.WX2,ZX.WX3,ZX.WX4,ZX.WX5,FR.FRONT_NUMBER,FR.WOF
```

```
A,FR.WOFD,FR.TIME_MIN,FR.TIME_MAX
FROM BATTBELLIACTWE_JOIN ZX,FRONT_WIDTHS FR
WHERE ZX.ISQNO=FR.ISQNO;

CREATE OR REPLACE VIEW BATTBELLIACTWEFRONT_JOIN AS
SELECT
ZX.ISQNO,ZX.WAR,ZX.NAME,ZX.LOCN,ZX.POSTYPE,ZX.POST1,ZX.POST2,ZX.FRONT,ZX.DEPTH,ZX.T
IME,ZX.AEROA,ZX.SURPA,ZX.CEA,ZX.LEADA,ZX.TRNGA,ZX.MORALA,ZX.LOGSA,ZX.MOMNTA,ZX.IN
TELA,ZX.TECHA,ZX.INITA,ZX.QUALA,ZX.RESA,ZX.MOBILA,ZX.AIRA,ZX.FPREPA,ZX.WXA,ZX.TERRA,ZX
.LEADAA,ZX.PLANA,ZX.SURPAA,ZX.MANA,ZX.LOGSAA,ZX.FORTSA,ZX.DEEPA,ZX.WAR_INITIATOR,Z
X.WINA,ZX.PARENT,ZX.ATTACKER,ZX.NAM,ZX.CO,ZX.CODE,ZX.CAV,ZX.TANK,ZX.LT,ZX.MBT,ZX.ART
Y,ZX.FLY,ZX.PRI1,ZX.PRI2,ZX.PRI3,ZX.SEC1,ZX.SEC2,ZX.SEC3,ZX.STRPL,ZX.STRMI,ZX.ACTORS,ZX.N,Z
X.ACTOR,ZX.STR,ZX.WXNO,ZX.WX1,ZX.WX2,ZX.WX3,ZX.WX4,ZX.WX5,FR.FRONT_NUMBER,FR.WOF
A,FR.WOFD,FR.TIME_MIN,FR.TIME_MAX
FROM BATTBELLIACTWE_JOIN ZX,FRONT_WIDTHS FR
WHERE ZX.ISQNO=FR.ISQNO;


CREATE OR REPLACE VIEW BATTBELLIACTWEFRONTTR_JOIN AS
SELECT
NM.ISQNO,NM.WAR,NM.NAME,NM.LOCN,NM.POSTYPE,NM.POST1,NM.POST2,NM.FRONT,NM.D
EPTH,NM.TIME,NM.AEROA,NM.SURPA,NM.CEA,NM.LEADA,NM.TRNGA,NM.MORALA,NM.LOGSA,
NM.MOMNTA,NM.INTELA,NM.TECHA,NM.INITA,NM.QUALA,NM.RESA,NM.MOBILA,NM.AIRA,NM.
FPREPA,NM.WXA,NM.TERRA,NM.LEADAA,NM.PLANA,NM.SURPAA,NM.MANA,NM.LOGSAA,NM.F
ORTSA,NM.DEEPA,NM.WAR_INITIATOR,NM.WINA,NM.PARENT,NM.ATTACKER,NM.NAM,NM.CO,
NM.CODE,NM.CAV,NM.TANK,NM.LT,NM.MBT,NM.ARTY,NM.FLY,NM.PRI1,NM.PRI2,NM.PRI3,NM.S
EC1,NM.SEC2,NM.SEC3,NM.STRPL,NM.STRMI,NM.ACTORS,NM.N,NM.ACTOR,NM.STR,NM.WXNO,
NM.WX1,NM.WX2,NM.WX3,NM.WX4,NM.WX5,NM.FRONT_NUMBER,NM.WOFA,NM.WOFD,NM.TI
ME_MIN,NM.TIME_MAX,TR.TERRANO,TR.TERRA1,TR.TERRA2,TR.TERRA3
FROM BATTBELLIACTWEFRONT_JOIN NM,TERRAIN TR
WHERE NM.ISQNO=TR.ISQNO;


###############################END############################################
```