We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists



122,000

135M



Our authors are among the

TOP 1%





WEB OF SCIENCE

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us? Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected. For more information visit www.intechopen.com



Data Warehouse and the Deployment of Data Mining Process to Make Decision for Leishmaniasis in Marrakech City

Habiba Mejhed, Samia Boussaa and Nour el houda Mejhed ¹Université Cadi Ayyad, Marrakech Département Génie Informatique, Laboratoire Micro Informatique Systèmes Embarques Systèmes sur Puce. Ecole Nationale des sciences appliquees, ²Université CadiAyyad, Marrakech Laboratoire d'Ecologie et Environnement, Faculté des Sciences Semlalia, ³&Université Louis Pasteur-Strasbourg I, Laboratoire de Parasitologie, Faculté de Pharmacie, ⁴Universite Sidi Mohammed Ben Abdallah, Fes Département Génie Informatique, Ecole Nationale des sciences appliquées de Fes. ^{1,2,4}Morocco ³France

1. Introduction

In the last decade, the epidemiology applied more and more tools to help to make decision. The aim is to translate epidemic data using the modelling concepts of health information systems for the decision-making. The information is a value-increasing necessary to plan and control the activities of an organism with effectively. It is the raw material that will be transformed by information systems. Often, the availability of data makes it very difficult, if not impossible, to extrapolate the information that really matter. It is essential to have rapid and complete information needed for the decision-making process: the strategic indicators are extrapolated mainly operational data in a database, through a selection process or synthetic gradually. The widespread use of data analysis techniques has made the information system a strategic element and policy framework for achieving the business.

Therefore, the decision-making systems have emerged in the 80s (decision support system) and offer techniques and means to extract information from a set of memorized data. As a result, the volume of information collected during an epidemiological case study enables the development of new observing systems to analyze and extract some indicators as appropriate clinical decision and public health.

The clinical decision support provided epidemiologist technologies necessary to facilitate this difficult task (Degoulet & Fieschi, 1998), (Gilbert, 2004), (Teh, 2009).

The data warehouse remains a valuable tool for storage and data accessibility, it is defined as a collection of information that integrates and reorganize the data from a variety of sources and make them available for analysis and assessment to scheduling and decision making.

If the data warehouse used to store historical data, with the finality analysis, the data mining is defined as a process of exploration and modelling data in order to discover new correlations, to find trends and stable patterns in the data. It proposes a number of tools from different disciplines, in particular, to decision making in epidemiology [Daniel, 2005), [René & Gilles, 2001), (Pascal et all, 2007), (Stéphane, 2007), (Egmont et all, 2002), (Hang & Xiubin, 2008). Data mining combines between various sciences domains (Databases, Statistics, Artificial Intelligence) to construct models from the data, and under the criteria fixed in advance and make a maximum of knowledge useful to make decision.

In Morocco, leishmaniasis remains a severe public health problem. Many foci were described in rural areas of Ouarzazate (Rioux et all 1986), Essaouira (Pratlong et all, 1991), Azilal (Rhajaoui et all, 2004), Chichaoua (Guernaoui et all 2005) and Al Haouz (Boussaa et all, 2009), (Rioux et all, 1986) but also in suburban areas, in Taza (Guessous et all, 1997) and Fez (Rhajaoui et all, 2004) Marrakesh is an interesting study site because it lies close to the focus of cutaneous leishmaniasis in the south of Morocco (Ouarzazate, Chichaoua and Al Haouz), and current studies have classified the area of Marrakesh as being at risk of cutaneous leishmaniasis (Boussaa et all, 2005), (Boussaa et all, 2007).

As the best choice of a vector-control strategy is dictated by sandfly ecology, we try to simplify this complex of diseases and quantify the climatic factors which can determine the distribution and activity of sandflies vectors in Marrakesh city. According to the WHO (2005), the activity of sandfly fauna is affected by many climatic factors as temperature, humidity and wind, besides seasons and according to sandfly species.

In this chapter we present a simple contribution to the fight against leishmaniasis in Morocco. The idea is to propose to epidemiologists an application based on tools of data warehouse and data mining to help them to make decision. In the first time, we conceived and modelled our information system to establish the pattern of the database on which we are going to work. Then, we develop a data warehouse to store and extrapolate data collected in Marrakech city, and we used a data mining tools for Leishmaniasis data analysis to get a better decision-making.

We studied three forms of leishmaniasis according to sandfly vectors collected in Marrakech city by (Boussaa et all, 2005), (Boussaa et all, 2007): Phlebotomus papatasi proven vector of zoonotic cutaneous leishmaniasis with Rodents as a reservoir; P. sergenti proven vector of anthroponotic cutaneous leishmaniasis and P. longicusis potential vector of visceral leishmaniasis with canine as reservoir hosts.

2. Material and methods

A. Sandfly collections

Specimens were collected in Marrakech city (31°36'N, 8°02'W, 471m a.s.l.) between October 2002 and September 2003 as described by (Boussaa et all, 2005). The specimens caught were preserved in 70% ethanol. They were cleared in potash 20% and Marc-Andre solution, then dehydrated and mounted in Canada balsam (Abonnenc, 1972). The identification was made by examining the morphology of male genitalia, female spermathecae and pharynges. For Larroussius species, we revised our specimens according to results of [Boussaa et all, 2008).

B. Data analysis

We have a data file.xls containing all the information on the activity of sandflies P.Papatasi, P. Sergenti and P. Longicuspis based on climate change (date, temperature and density). We can present data from Excel files in the following diagrams:



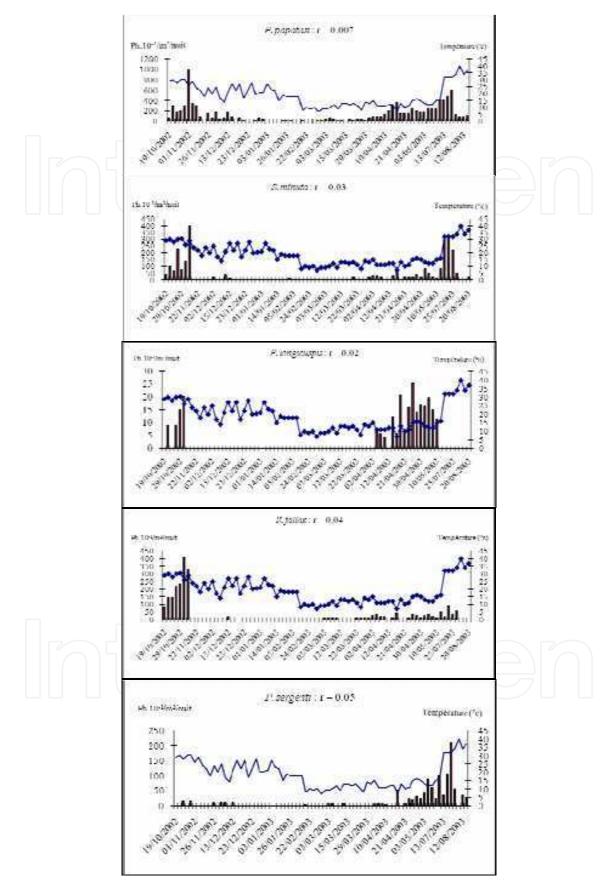


Fig. 1. The activity of sandfly population in Marrakech area

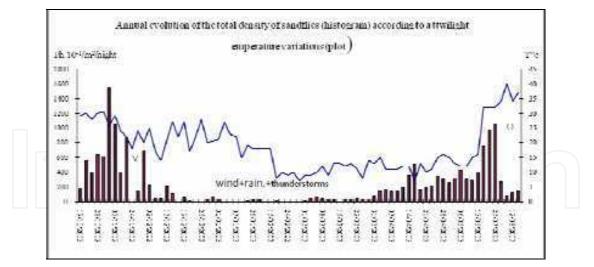


Fig. 2. Annuel evolution of the global density of the sandflies vector according to the time and the climatic change

3. Development of a data warehouse Leishmaniasis

In this section we proceed to the conception of our data scheme. Our investigations on the sandflies collections were carried out in Marrakech city. The goal is to design a database, to develop DW and to load this database, and then the relationships in the cube are built automatically to give the answer to question posed in this case.

3.1 Background

The creation of a data warehouse involves several steps:

The conception: the implementation of a data warehouse usually begins by framing the project, define the needs and goals expressed by policymakers, modeling and designing a data structure. There are two data models, the star pattern, in this model, we must define one (or more) table (s) made with one or several measures (values of indicators). Both must have multiple dimension tables whose primary keys form the primary key tables done. Warning: The dimension tables are not linked.

Then the model snowflake which is derived from the star schema where the tables are standard size (of the table remains unchanged). With this scheme, each dimension is divided according to his (or her) hierarchy (s).

The acquisition of data: The data will be extracted from the sources.

- a. The static extraction will be performed when the DW must be loaded for the first time and is conceptually a copy of operational data.
- b. The incremental extraction, is used for the periodic updating of the DW, and captures just the changes in data sources at the last extraction.

The choice of extracting data is based mainly on their quality, selection of data from the database is not a simple task to do.

Data cleaning: This phase will improve the quality of duplicate data, inconsistencies between the values logically related, missing data, unexpected use of a field, impossible value or wrong ...

Loading DW: The loading of data in the DW is the process is to load the data cleaned and prepared in the DW.

3.2 Needs specification

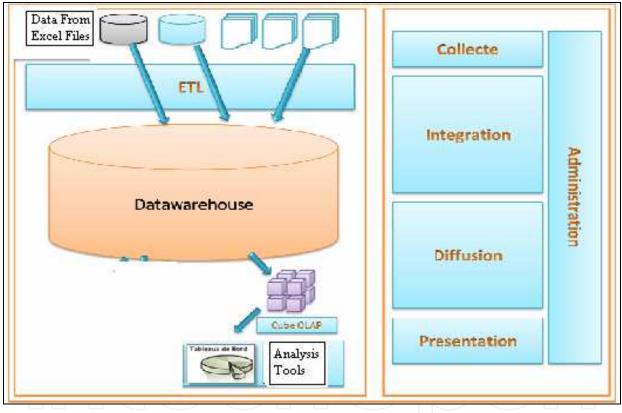
We have developed the analysis tools concerning, the population of sandflies, according to the various species listed in Marrakech city and their density, we considered the human population which may become affected after bites of infected sandflies.

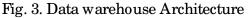
These tools allow knowing the following information: (a) The density of each species of sandflies listed, (b) The period at risk for the spread of the disease, (c) The rate of infection of humans by infected sandflies, (d) The rate per unit time, which a man loses her immunity and becomes susceptible, (e) The rate of infected and susceptible in humans.

To meet the needs of decision makers, we implemented the DW with respect to the architecture described in the following section.

3.3 Data warehouse architecture

Schematic below shows the architecture of the data warehouse applied to Leishmaniasis data.





To ensure a robust, flexible and portable solution, we adopted a software architecture divided into several parts:

Collection of raw data files: the job of this module is periodically connected to all servers, to check the generation of new data files

Conversion of the Files: The application of this module is written in Java. This module convert the data files collected by the module above, and filter the information contained therein. It permits to leave only the information that will be used in the future treatments.

Loading data into the data warehouse: ETL process allows Extraction, Transformation and Loading of data from various sources (databases, files) into DW. ETL process is the most important module to design a Data warehouse with respecting two constraints: data sources and data types (data quality).

Building of multidimensionnel cubes :

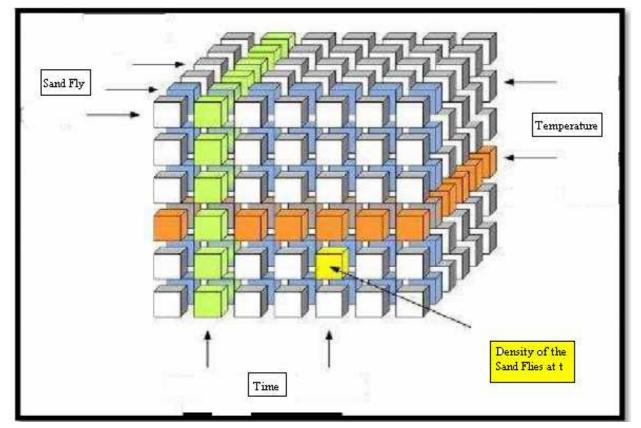


Fig. 4. Example of multidimensional view (Cube) of the Leishmaniasis data.

Creation of the graphical interfaces, graph and report: used for operation, querying cubes and creating reports.

3.4 Data model

This section deals with the transmission of the Leishmaniasis disease from the sandflies vector to human. There are four actors in this case, sandflies species, climatic change and Time and Human. The data dictionary given in the following table:

data	Description of data	
Human	Sexe	
	Age	
	statut	
sandflies	specie	
	Name	
Temperature	Degree	
time	wear	
	Month	
	Day	

The data layer architecture of Leishmaniasis is illustrated schematically by:

Data Warehouse and the Deployment of Data Mining Process to Make Decision for Leishmaniasis in Marrakech City

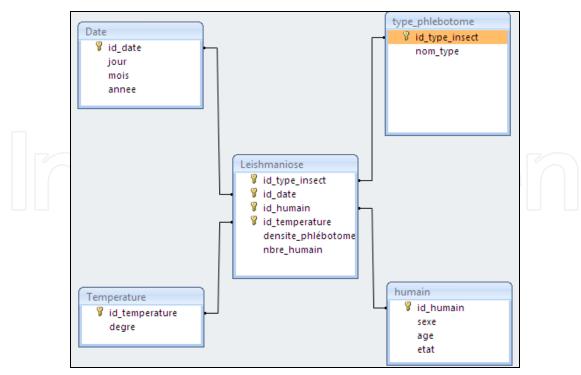


Fig. 5. The model data warehouse diagram

The model chosen must comply with the requirements and needs of use, in our case, we opted for a star pattern respecting the nature of the information we have.

Dimension Tables: Date, Temperature, Sandflies (P. Sergenti, S. Minuta, S. Fallax, P. Logicuspis, P. Papatasi), Human.

Fact Table: Leishmaniose.

Data is extracted from Excel files using java code. The program consists of two classes, one for extraction and one for export useful data to the database.

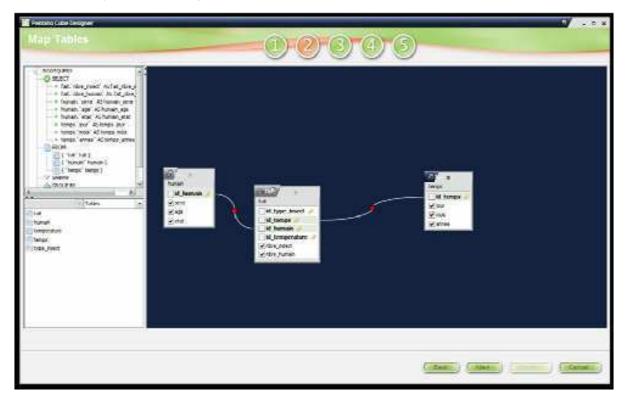
Permate Units Same Tanine Common Tanka Inter Tanine Endore Endore Reporting Freement: Endore Endore Reporting Inter Tanine Endore Endore Reporting Freement: Endore Endore Reporting Inter Tanine Endore Reporting Inter Tanine Endore Endore Reporting Inter Tanine Endore Endore Reporting Inter Tanine	opentaho	Serer Statur: O' 6 4			22.5
Advancestable Interface Advancestable Interface Conditionmen Interface	disc.	THEY'R RATE AND DOOR DOOR DO	ense [Scheker]		
Conductories Conductories <th>Abrenetator</th> <th>Sets Section - O TempinCata</th> <th></th> <th>and the second se</th> <th></th>	Abrenetator	Sets Section - O TempinCata		and the second se	
Nement Learning out plot Down Permit N 0000 Strangeristic Lear transi Permit N 0000 Strangeristic Lear transi Permit N 0000 Strangeristic Institution Permit N 0000 Strangeristic Permit N 0000 Strangeristic	Conducione				
Comment Links Comment Links Comment Links Comment Comm	and a second sec	1			
Communit Lipida Inclus: Lipidas hepototos Neuto: Lipidas hepototos Venu: Ecos Venu: Ecos Venu: Ecos Venu: Ecos Venu: Ecos Venu: Como Venu: Co	Personal of Colda States (State		(inertienel)		
Review Linkins Reportory Review Linkins III Second View Rome New Roberts Data Storms Consolitions United as Enterprise Linkins Age					
View Ences New R Backet Memory Control States Chronologie Vandet Enterprise Editors Sea	Recture Londont Reporting Restart, Factories III, Second		Present 2		
(dic nyng Hisahof lash	View Brans View BLIADAC INCLUS Check Data Saurus, Chromotona		unu Internet Reported		

3.5 Dimensional cube

Pentaho integr a CubeDesign tool, it allows to have the cube in XML format and to publish it in the User Pentaho Console. Cube creation through 5 steps as shown in the below: Step 1. With the CubeDesigner we establish the connection to database Leishmaniose.

Persona Cele Tradicion	2
Sturt Rere	-00000
Table scalar and description. Sheaty the neme and description for the case. The case neme will be cased for the o	e substituizante las est antidas esticas mont
Cate Lane	
15 Norme of Belger name for longer	Million III III III IIII IIII IIII IIIIIIIIII
	phonese for a second se
rdarinak Granovaria Sarat Sarat Sarat Rasi	Feet Consider as sociedal
	(and (date) (construction)

Step 2. Select tables and views useful to visualise the Cube. For example, to calculate the density of the P. Sergenti and the infected human in a fixed interval time.



Pentalo Caba Designer						J
Create Measures		-0	00	4)	9	
Create Neuranes. Next the Set take that contact the Gale record	ons they change the	recourse catalons and gardy t	In aggright further, fait		19 1970.	
Fact Table	1.64					
N:	(*)					
Fact Table Columns		Measures				
(allow real	11	Report Relat	Surgeie .	Acres	Steary Gener	
10002000		Tel: Regulater	828	1,778	stand building turners	
		13.22.02.0				
	10 20					
	Contract of the					
	Committee of					
		1				
1				_		
					CONSI- Company	(Kata)
					Contraction of the second	

Step 3. Choose of the measures: Infected human

Step 4. All the dimensions and their aggregation will be specified via the interface below.

and a set of the				5
Create Dimensio	ons in the second	-00	390	
Create Discretario Create Discretario Create Stransform from the educe the Sottom section. Cox "App in	a (NG ang "Ali Neo Stanato" an Neo "Juttor ti ali neolar anget	con. You can add lenab by cangitha Anon but All fy a lengt	ora. Once dimensions are created, elect the tim	enelineelijensta na steraja ta atstala a
(All fam Drowners) Servic fields		Secolars		
Nacione pone 1 Nacione pone 1 Nacione pone Nacione pone Nacione Nac	(come	n 25 jans units see n 25 jans n 25 jans of the program of		
Attributes and Properties		One-rank for the Sharesen		
Calcol Rocky (Calcol Rocky) (Calcol	Resource Resource of American America American American Americ American American Americ Ameri		S Inc.	
			(1100)	(100) (100)

Step 5. The interface below allows us to view the file for the model; it's possible to update it as needed.

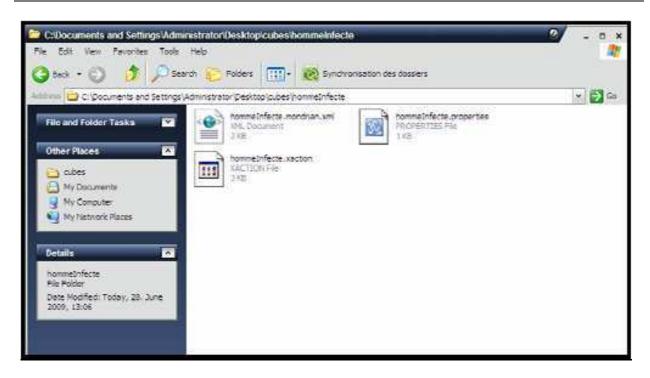
Central Cale Teager		11
Final steps	00000	
Ship 1: Review the XML that the Yeye XML factors the scheme lefts	ter for sear rates	
Constant Sing 2: Profession Sing 2: Providen the codes Constant Constant	Contrast space - Space (C) Contrast space - Space (C)	

Finally, the pattern is saved

Constant and a second second		9
Final steps	00000	
Nay 1: Sovice the 208 Do by Yes MC patients reachers period	Minur cos.	
(Internet)		
step is midde. Do the Note' Setter to public microstic scheme of	dine 🤺	*
(Sive Information	
Step 3. Proview the cate Dol the Therein' buttor to premier the cote samp th	an under C. Die neuer tetres kernen en Dielen nies	
	(anna) (annali)	
Shap 4: Save City Shaff Saves to as your your you to a local Se.		
	1 <u>+0000</u>	

Tree files are generated : (a) Xml file to save the pattern produced by CubeDesigner. (b) Properties file : for the allocation of the database. (c) Xaction file : presents a set of all protocols to data access.

Data Warehouse and the Deployment of Data Mining Process to Make Decision for Leishmaniasis in Marrakech City



3.6 Pattern publication with workbench

Workben is a tool to create diagrams, for our case it is just used to refine and publish the pattern designed by CubeDesigner.

Step 1. Configuration of the Workbench references to establish connection to database.

Workbench Preferences	2	<u> </u>	3
Niver Class Name	com mysql jabe Driver		15
Connection URL	utto: mysqLiflocamostileish		1
lser name	hout		
assword	E.		15
Schema (Optional)	1		1
Require Schema Attributes	C		
		Accept Cance	1

www.intechopen.com

183

Step 2. Pattern raffinement.

B-B-B-E			
C feiferen immerialitie pare	and an an and	N M M S & X N S B W	<i>id</i> (
B Direce B Direce B Direce B Low C So C So	A Biology Terry Record Collection of the second	Interest and and and an an and an and an and an an an and an an and an	

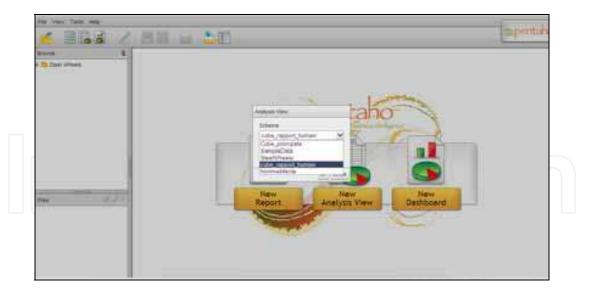
Step 3. Now, it is important to have URL for the host user, The password for the publisher and The login and password of the user server.

5-058 E			
C Schenne - heteretetteche die			a*15 15
8 1 × 10 M	nia di 🖉 🖽 🖉 🎦	S 2 X 5 0 1 D	
 Briteriu Commentation Solid A fame 	+) Bit Base Server	Historichy for Temps' Desentance	
+ A ter - A ter S substances ter	etanture studiender namenfesterung presenterfass	Security ERE Processing Product 100000 percentation Product Of Decement of	
		Peodate Caluminas Uniter Information Permanent	
		2 Remains the Software	
·			

3.7 Visualization phase

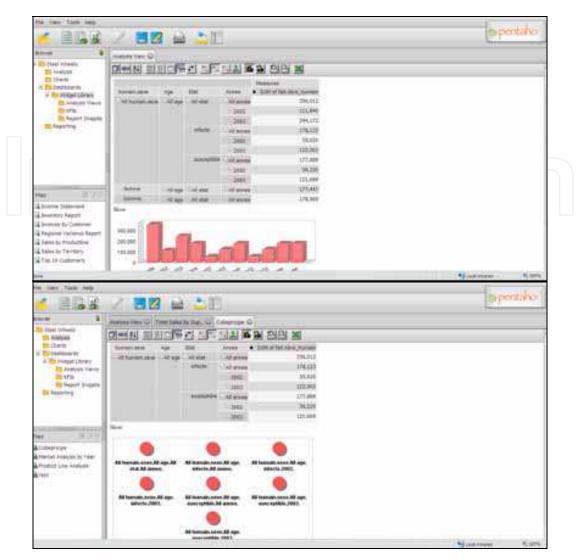
'User Console' tool gives different views of the pattern previously published.

Data Warehouse and the Deployment of Data Mining Process to Make Decision for Leishmaniasis in Marrakech City



It's possible to have several modes to visualize data from database. As illustrated in the following figures.

6	1 32	100 T		tapenta
Rent I		e Sentrite, Q. Cou		
a dia Meri Liberte				
D REPORT	Chart Properties	1471 19 - 111		
in the law income to	Chail Trpe	In Costs by Course		
at Consiger Sillines	Toutte Dell Decemb	Vehice Be		
#1 (P3	Chet Stie	Venca Bar 30 -	the second se	
Report Signat	Chart Enter Front	Herbarda Bar 30 Discher Verbar De	1.18.00	
	Vicewatal aris blut.	Dischart Verbrat Bar 301	Line.	
	Verigent mit lebet	Dischart Holgostal Bar Discheri Horgostal Bar 30	and the second s	
	1.2.2.2.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1	Vehc# Lns Historial Lise	and the second se	
PPPE SECTO	Airy Label Frank	Vallog And	(2) #	
a counter	Asis Tak Label Bat	Vellog Stated Are	U = 3 =	
Tertal Acolumn by Tear	Mere Legend	Personal Statistics		
American Constant	Lapsed Fort	(Pre Charle by April	3.4.6	
	them 54cm	a particular in	a contract of the second se	
	Tarre Free	Setslarf . A Plac	(m) (12.m)	
	Clart Height	210 Cher Nubb	1940	
	Sexprod 3.0.8	100 (120) (20)		
			ES CHIEF	
tia view Tann iven			ES (Secol	-guarran
In view Tast. wes South Strange South Strange So				e juerree Opentado



The hearlth professional may use OLAP concept to fialter and visualize other type of information.

			-				
- Hereite		A DESCRIPTION OF A DESC					
Es distanti forma		TO MEN BUILDER	対応開	THAT	14 신신 제		
IS Charle		Column Q					
· · ·	pet strang-	Contraction of the local division of the loc					
the second	navail travel	N Rooms					
	100	a la contra c					
	mand Josephere						
(D Autor	- C	210kar					
		Oliferry					
		Des Liveren					
	10.00	housing by	1	Arres	· Data of the store because		
	- Status	ad human person () of sign (to and				
and it is a second s							
A Contract of the	in he has		offecte-	All Arrest			
a Contra cont A Mariles Anal A Printer Line			offecte-	- 2016	15,428		
Water ing			(Palla	- 2010	15,428 122,849		
Alteriation Anna Alteriations			1040	- 2016	15,425 102,649 1,17,849		
Alteriation Anna Alteriations				- 2010	15.425 102.445 1/7.445 96.137		
A Harbert Anna A Headard Line				2011 - 2011 - 34 Annos	15,425 102,649 1,17,849		
A Harbert Anna A Headard Line		la.		1992 1992 1992 10	15.425 102.445 1/7.445 96.137		

186

4. Data mining: application to the Leishmaniasis

Given the seriousness of leishmaniasis in Morocco, it was essential to deploy easy methods to reduce its exploitable spread if not eradicate it completely. Our proposal aims to exploit tools of data mining process on this infectiou disease. By definition, the data mining attempts to extract knowledge from vast volumes of data.

The wealth of information transmitted on vectors of disease allows us to apply these tools to identify methods and anticipate behaviour, therefore, make a good decision.

4.1 K-means and deployment

The technologies that are on the market, offers comprehensive platforms and integrated data analysis to meet all requests of indicators developed in the industry. We are able to accede to any type of data stored in our data base, to implement operations to analyze the data and present results in a need predefined by the user.

The software that we have developed our application offers a wide range of approaches ranging from methods of descriptive statistical analysis to predictive modeling methods.

The first step is to create a new diagram and import the data as shown in the screenshot below.



4.2 Descriptive statistics

We can do descriptive statistics to variables. We calculate the frequency histograms on all columns to count the number of active and additional comments.

a and the second s						
. besette	-			Theory of Lot of	00	
 B (particular and) B (particular and) 	AND MARKET 11 Recording 117			AD ADDRESS		
	-	And and a second se				
	and a second section of the		Sand Street	And a local division of the local division o		
		N 10 M 1	- 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	The Third		
		DOM: NOT	1			
		pt-m-146	1			
		A-9-200	1			
		an an and	1 5.84			
		PU11100E	1 (Law			
		MPTCOME.	10 140			
		1011/046	10 1000			
		an recent	A			
		34-11-1866 34-11-1865	1 540			
		#1-1112#18				
			1 5,00			
		an Island	- 1,81			
		or the same	i Las			
		STATISTICS.	1			
		ministery.	1 3			
		Printed				
		and the second se	1 1 1 1 1 1 1			

4.3 Method principal component analysis (ACP)

They are three categories of data mining algorithms: supervised methods, unsupervised methods and methods of data reduction. Each category is based on a number of techniques. In this section, we chose the third type using the method of principal component analysis (ACP).

Given a set of observations described by variables exclusively digital (x1, x2, ..., xp), the APC aims to describe the same data set with new variables in reduced numbers. These new variables will be linear combinations of original variables. Principal component analysis can therefore be seen as a technique to reduce dimensionality.

4.4 Visualization of our data

To implement the ACP method, we can see, for example, date and temperature data concerning P. longicuspis. After we define an analysis of the variables studied. The result is given in the following figure:

. Trive the	A CONTRACT OF A
 B Server (Find) B Server (Find) M Hermon Description (Matching) 	Name -
	Server data (and) the C
	Annual citizen and the second s
	Elgen values
	Multile Read + 1 100
	Test light same Surgentiant Miningram Surgentiation
	4 Later water and the second sec
	MA CAREER CONTRACTOR CONTRACTOR
	Factor Loadings [Communality Estimates]
	Advanta Aug. A
	104 0.04 M
	A financial and the second sec
	The Asian Distances International State
	the data
	The second se
	Eigen vectors – Factor Scores
	All the second s
	A sugarda 22000 (K2000 - Later)
	Casadadat has inter-
	Consider the Colline State (1)
	in the second seco
the multiplice (Reserve)	Republican optimis interfere ferror companies factories factories for the factories for the factories of the
Claiming Inchange	Reformant tooling Backwards Backwards Backwards

To better assess the relative positions to sandflies in the first factorial design, we add the component display. We put abscissa variable representing the first axis, calculated using the ACP, and ordered the second axis. We get the point cloud :

E G. Telation	PROJECT 14	-
 E. Denote (PCris) C. Define (educ) ? S. Define (educ) ? S. Define (educ) ? L. Destroyed Y. 		_
	4 4 4	

Fig. 6. Point cloud data on (date, temperature, for P. longicuspis)

Among the variables are, we want to check the effect of the variable date that can distort our results. We will color the points according to this variable, we select as a variable component as is shown in this illustrative visualization.

Dapas Company banks and	
Same (Print)	PALANCE - P
Set Provided Longovich Burgari 1 () ()	

Fig. 7. Using the date_format as illustrative variable

Now we apply the ACP method on all data:

Twind its	New York Concerning Street Str
 Setsual (downed/FLDH) Children (name) 1 Set Principal Component tradition 1 	Martin and Statement (Conception)
In Interested (

Fig. 8. Implementation of the ACP method on sandflies data

The most difficult decision in any project is to determine what method should be implemented. In our case, we prefer use the open source tools:

- a. ETL « Talend Open Studio » Talend Open Studio offer: A complete range of components, Traces and statistics of treatment in real time, the addition of specific code and the integration on the decisional open source.
- b. MySQL to create the database and data warehouse for the persistent storage area.
- c. Pentaho BI Suite: cover many areas of Business intelligence through various software (owned by Pentaho or integrated into).
- d. Pentaho Analysis (Mondrian + Pivot), Cube Designe, schema Workbench. For analysis OLAP
- e. Open source data mining software Tanagra.

5. Discussion

Our investigations were conducted in Akioud, an urban district in Marrakech city, during one-year-study. This site was selected, considering the presence, of all the sandfly species inventoried in the urban area of Marrakech (Boussaa et all, 2007).

According to the correlation between the weekly density of the three vectors (P. papatasi, P. sergenti and P. longicuspis) and the factor R0, we can prevent the risk of leishmaniasis in this area.

- a. For P. papatasi population, R0 factor is superior to 1 during two periods of the year: November and May-June-July, which correspond to the periods of risk of zoonotic cutanous leishmaniais caused by L. major in this area.
- b. For P. sergenti population, R0 is superior to 1 during the period of July–August and inferior to 1 in the rest of the year. So, this period corresponds to the phase of risk of anthroponotic cutanous leishmaniais caused by L. tropica in this area. We observe that R0 reaches its peak during the period of August when the temperature is very high.
- c. For P. longicuspis, the results have shown two periods of risk of visceral leishmaniasis in this area: October and May-June.

(Boussaa et all, 2005) classified Marrakech area as being at risk of cutaneous leishmaniasis because of the high density of P. papatasi throughout the year, its position close to the cutaneous leishmaniasis foci in the arid region (Rioux et all, 1986) and the omnipresence of Meriones shawi, main L. major reservoir host in Morocco. In this work, we approve the conclusion of [Boussaa et all, 2005) and we exploit these data to assist in the decision on the issue of the fight against leishmaniasis in Morocco.

Indeed, these are spatio-temporal models which permit to follow and control the evolution of leishmaniasis in terms of time and region, and to find the appropriate threshold of the population of sandflies for stopping the multiplication of the disease. In Chichaoua (70 Km from Marrakech city), focus of anthroponotic cutanous leishmaniais, (Bacaër and Guernaoui, 2006) suggest that the epidemic could be stopped if the vector population were reduced by a factor (R0)2 = 3.76.

6. Conclusion

To fight against the spread of the cutaneous and visceral leishmaniasis and to address the need of the responsible for population health to set the policies that determine the nature of health care provided, and most importantly, for fast evaluation of the severity of the multiplication of the disease, we proposed to develop a model of data warehouse to storage appropriate data of the sandflies seasonality and the information about the susceptible human (patient with suspected Leishmania infection). Then, structural key and functional parameters will be measured to ensure the advanced treatment. By employing data mining process we have the ability to extract knowledge and to generate summaries for better health decision-making. Through this complete and operational platform, it will be easy to control the sandflies seasonality and the rate of the transmission disease.

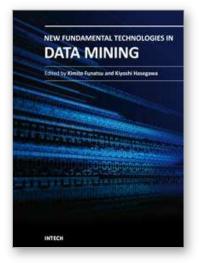
7. References

- P. Degoulet et M. Fieschi (1998). Informatique et santé Collection, Paris, Springer-Verlag, Volume 10,
- Gilbert Saporta (2004). Data mining ou fouille de données, RST « Epidémiologie » *Data Mining*.
- Daniel T. Larose (adaptation française T. Vallaud) (2005). Des données à la connaissance : Une introduction au data-mining (1Cédérom), *Vuibert*.
- René Lefébure et Gilles Venturi (2001). Data Mining : Gestion de la relation client, personnalisations de site web, *Eyrolles*, mars 2001.
- Pascal Poncelet, Florent Masseglia and Maguelonne Teisseire (Editors) (2007). Data Mining Patterns: New Methods and Applications, Information Science Reference, ISBN: 978-1599041629, October 2007.
- Stéphane Tufféry (2007). Data Mining et Statistique Décisionnelle, Technip, nouvelle édition revue et enrichie, juin 2007.
- Egmont-Petersen, M., de Ridder, D., Handels, H. (2002). Image processing with neural networks -a review. *Pattern Recognition* 35: pp. 2279–2301. doi: 10.1016/S0031-3203(01)00178-9. 2002.
- Boussaa, S., Guernaoui, S., Pesson, B., Boumezzough, A. (2005). Seasonal fluctuations of phlebotomine sand fly populations (Diptera: Psychodidae) in the urban area of Marrakech, Morocco. pp. 86–91, *Acta Trop.* 95.

Boussaa, S., Pesson, B., Boumezzough, A. (2007). Phlebotomine sandflies (Diptera: Psychodidae) of Marrakech city, Morocco. pp. 715-724, Ann. Trop. Med. Parasitol. 101.

WHO (2005). Lutte contre les leishmanioses. Série de Rapports Techniques.

- Rhajaoui, M., Fellah, H., Pratlong, F., Dedet, J.P., Lyagoubi, M. (2004). Leishmaniasis due to Leishmania tropica MON-102 in a new Moroccan focus. *Trans. R. Soc. Trop. Med. Hyg.* 98, pp. 299–301.
- Teh and All (2009). Development of a Data warehouse for Lymphoma Cancer Diagnosis and Treatment Decision Support. *Proceedings of the 10th WSEAS International Conference on MATHEMATICS and COMPUTERS in BIOLOGY and CHEMISTRY*, Pp. 15 _24. ISSN: 1790-5125, ISBN: 978-960-474-062-8.
- Abonnenc E. (1972). Les phlébotomes de la région éthiopienne (Diptera: Phlebotomidae). *Mémoire de l'ORSTOM*. 55, pp. 1–289.
- Bacaër, N., Guernaoui, S. (2006). The epidemic threshold of a simple seasonal model of cutaneous leishmaniasis. *J Math. Biol.* 53, pp. 421–436.
- Boussaa, S., Boumezzough, A., Remy, P. E., Glasser, N., Pesson, B. (2008). Morphological and isoenzymatic differentiation of Phlebotomus perniciosus and Phlebotomus longicuspis (Diptera: Psychodidae) in Southern Morocco. *Acta Trop.* 106, pp. 184– 189.
- Boussaa, S., Pesson, B., Boumezzough, A. (2009). Faunistic study of the sandflies (Diptera: Psychodidae) in an emerging focus of cutaneous leishmaniasis in Al Haouz province, Morocco. *Ann. Trop. Med. Parasitol.* 103, pp. 73-83.
- Guernaoui, S., Boumezzough, A., Pesson, B., Pichon, G. (2005). Entomological investigations in Chichaoua: an emerging epidemic focus of cutaneous leishmaniasis in Morocco. J. Med. Entomol. 42, pp. 697–701.
- Guessous-Idrissi, N., Chiheb, S., Hamdani, A., Riyad, M., Bichichi, M., Hamdani, S., Krimech, A. (1997). Cutaneous leishmaniasis: an emerging epidemic focus of Leishmania tropica in north Morocco. *Trans. R. Soc. Trop. Med. Hyg.* 91, pp. 660–663.
- Pratlong, F., Rioux, J.A., Dereure, J., Mahjour, J., Gallego, M., Guilvard, E., Lanotte, G., Périères, J., Martini, A., Saddiki, A. (1991). Leishmania tropica au Maroc. IV. Diversité isozymique intrafocale. Ann. Parasitol. Hum. Comp. 66, pp. 100–104.
- Ramaoui, K., Guernaoui, S., Boumezzough, A. (2008). Entomological and epidemiological study of a new focus of cutaneous leishmaniasis in Morocco. *Parasitol.* Res. 103, pp. 859-863.
- Rioux, JA., et all (1986). Les leishmanioses cutanées du bassin méditerranéen occidental: de l'identification enzymatique à l'analyse éco-épidemiologique, l'exemple de trois 'foyers', tunisien, marocain et français. Montpellier, France: Institut Méditerranéen d'Etudes Épidémiologiques et Ecologiques. pp. pp.365–395.
- Hang Xaio, Xiubin Zhang (2008). Comparison Studies on Classification for Remote Sensing Image Based on Data Mining Method, WSEAS TRANSACTIONS on COMPUTERS. Volume 7, ISSN: 1109-2750, pp. 552 558.



New Fundamental Technologies in Data Mining Edited by Prof. Kimito Funatsu

ISBN 978-953-307-547-1 Hard cover, 584 pages **Publisher** InTech **Published online** 21, January, 2011 **Published in print edition** January, 2011

The progress of data mining technology and large public popularity establish a need for a comprehensive text on the subject. The series of books entitled by "Data Mining" address the need by presenting in-depth description of novel mining algorithms and many useful applications. In addition to understanding each section deeply, the two books present useful hints and strategies to solving problems in the following chapters. The contributing authors have highlighted many future research directions that will foster multi-disciplinary collaborations and hence will lead to significant development in the field of data mining.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Habiba Mejhed, Samia Boussaa and Nour el houda Mejhed (2011). Data Warehouse and the Deployment of Data Mining Process to Make Decision for Leishmaniasis in Marrakech City, New Fundamental Technologies in Data Mining, Prof. Kimito Funatsu (Ed.), ISBN: 978-953-307-547-1, InTech, Available from: http://www.intechopen.com/books/new-fundamental-technologies-in-data-mining/data-warehouse-and-the-deployment-of-data-mining-process-to-make-decision-for-leishmaniasis-in-marra

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri Slavka Krautzeka 83/A 51000 Rijeka, Croatia Phone: +385 (51) 770 447 Fax: +385 (51) 686 166 www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai No.65, Yan An Road (West), Shanghai, 200040, China 中国上海市延安西路65号上海国际贵都大饭店办公楼405单元 Phone: +86-21-62489820 Fax: +86-21-62489821 © 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the <u>Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License</u>, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.



