PRELIMINARY RESULTS OF THE ANALYSIS OF FIELD DATA FROM OVENS

Oliviu **MATEI**
Technical University of Cluj-Napoca, North University Centre of Baia Mare
oliviu.matei@cunbm.utcluj.ro

Keywords: ovens, statistics

Abstract: *The paper presents the preliminary results after analyzing the field data from 85 ovens deployed in the United States. We research especially the behavioral pattern of the user regarding the openings of the door and we prove that there are obvious relations between the pattern and some specific features, such as day of the week and hours of the day.*

1. INTRODUCTION

This paper presents the preliminary results after analyzing the data field collected from 85 ovens deployed in the United States.

We research a database of more than 1,000,000 records for several electrical appliances installed in real test field, containing data from ovens as well as from refrigerators. However our focus is on the behavioral pattern regarding the openings of the oven doors because there was no availability for refrigerator openings. For our specific purpose, we took into account a specific model of oven, namely Kenmore Elite 97102, having the serial number E02A82C5BEC0.

Of course, this is not the first attempt to analyze field data. This is crucial for any serious research and has been done in several domains, such as education [1], seismology [2], wildlife [3] and ecology [4].

On the other hand, ovens are part of our daily life, therefore a significant amount of scientific literature has been dedicated to them, especially with respect to temperature, home

and related aspects by Darigny et. al. [5], Omura et. al. [6], Li et. al. [7] and Boulet et. al. [8] and energy consumption [9].

2. METHODOLOGY

The data covers roughly 7 weeks: 15.11.2013 – 25.12.2013 and consists of 305 records. The data looks like in table 1.

Table 1. A snapshot of the database containing the data to be analysed

Timestamp	DW	Date	Hour	Door_Close	
15.11.2013 00:28	6	15.11.2013	0	TRUE	
15.11.2013 00:28	6	15.11.2013	0	FALSE	
15.11.2013 01:11	6	15.11.2013	1	TRUE	
15.11.2013 01:11	6	15.11.2013	1	FALSE	
15.11.2013 01:40	6	15.11.2013	1	TRUE	
15.11.2013 01:41	6	15.11.2013	1	FALSE	
15.11.2013 01:44	6	15.11.2013	1	TRUE	

The *Timestamp* and *Door_Close* are from the database. The day of the week (*DW*), *Date* and *Hour* fields have been drawn from the timestamp. A preliminary inspection of the data shows that the values are consistent (e.g. the number of records with *Door_Close = TRUE* is equal with the number of records with *Door_Close = FALSE*).

We research the correlations:

a) Day of the week vs. number of openings of the door of the oven;
b) Hours of the day vs. number of openings;
c) Days of the week vs. hours when the door was opened.

3. RESULTS

3. 1. Day of the week vs. number of openings of the door of the oven

The scatter plot derived from the data depicted in Fig. 1. The *X* axis represents the days of the week (1 = Monday… 7 = Sunday). The *Y* axis shows the number of openings.

The statistical correlation factor (as defined by Hedges [10] or Hotelling [11]) between the days of the week and the number of openings is 0.1097, which shows that there is no statistical correlation between the two. But that does not mean that the two data sets are not related. For instance, it is obvious that on Tuesday the oven is not used at all.

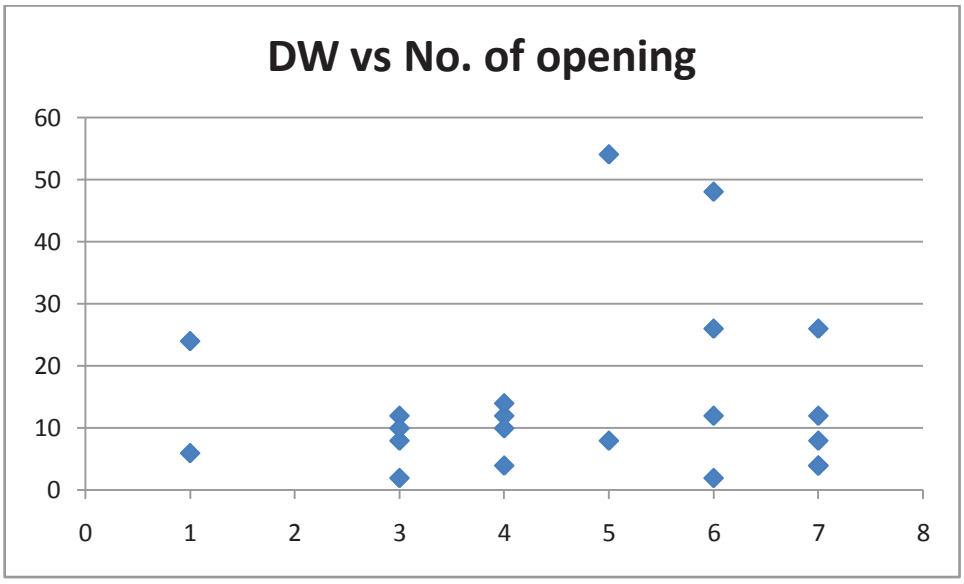

Fig. 1 – Day of the week vs. number of openings

3.2. Hours of the day vs. number of openings

The scatter plot is depicted in Fig. 2.

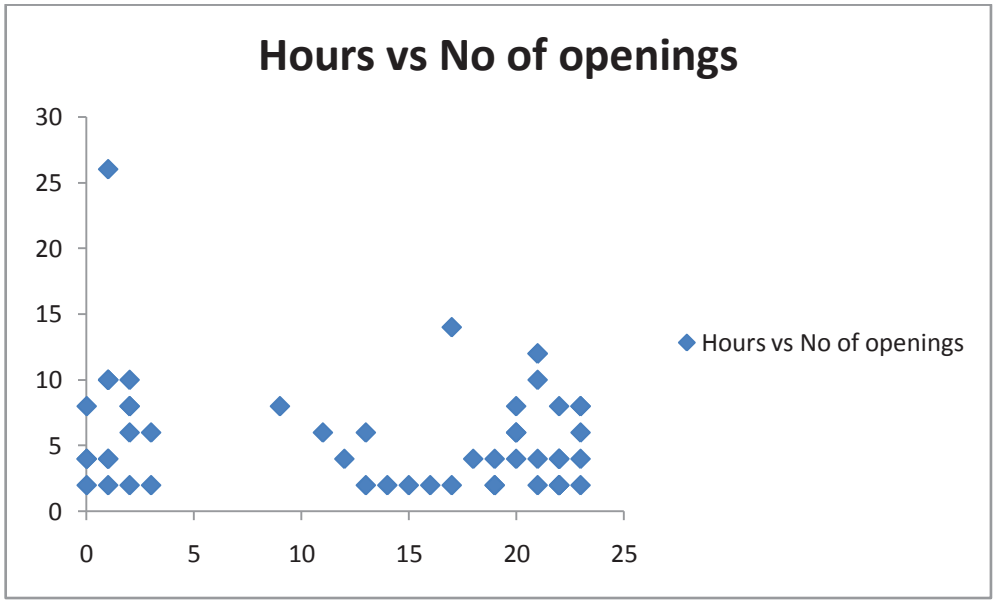

Fig. 2 – Hours of the day vs. number of openings

The *X* axis shows the hours (0..24) and the *Y* axis represents the number of openings. For relevance of the data, the time consists only of hours, not of minutes.

The correlation factor is -0.202 which says that there is no significant statistical correlation (although more than in the previous case – DW vs. number of openings). However, some conclusions can be drawn, such as that the oven is used over the night and is idle from 4:00 to 9:00.

3.3. Days of the week vs. hours when the door was opened

The scatter plot is shown in Fig. 3. The *X* axis represents the days of the week (1 – Monday,.. 7 = Sunday) and *Y* axis shows the hours when the door was opened. For relevance, the time contains only the hours, not the minutes. The correlation factor is 0.0247, which means that there is no statistical relation between the two data sets.

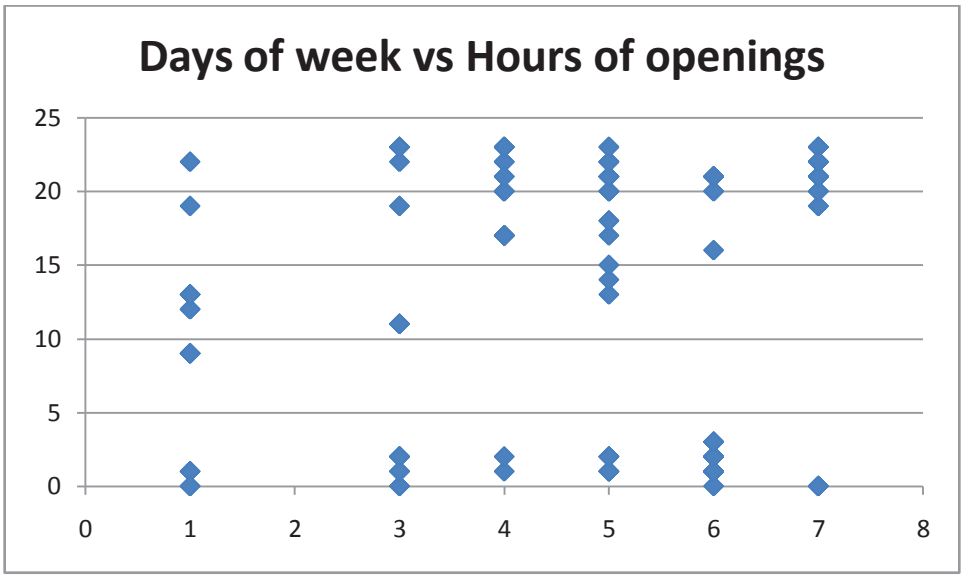

Fig. 3 – Days of the week vs. hours of the day when the door was opened

However, a close look at the plot shows that:

- The oven is not used on Tuesdays (the conclusion was draws also previously – see section 3.a));
- The over is not used from 4:00 – to 9:00 (see section 3.b));
- The time when the oven starts being used goes later over the week:
  - on Mondays, the oven starts working around 9:00;
  - on Tuesdays: off;
  - on Wednesdays: starts around 11:00;
  - on Thursdays: starts around 17:00;
  - on Fridays: starts around 13:00;
  - on Saturdays: starts around 16:00;
  - on Sundays: starts around 19:00.

4. CONCLUSIONS

In this research we analyzed only raw data, with no preprocessing (e.g. using only relevant data, by removing exceptions). However, the preliminary results are good and promising.

The first conclusion is that the data is consistent and relevant. Although there is no statistical correlation between the datasets compared, we could draw some interesting observations:

- The oven is not used on Tuesdays;
- The over is not used at all from 4:00 – to 9:00;
- The time when the oven starts being used goes later over the week.

The data processed covers only seven weeks, which is a reasonable interval, but to make sure that the observations are correct, they should be applied on larger intervals, at least at this research stage.

5. FURTHER RESEARCH

We used all the data, but we can remove some exceptional situations, which will make the observations more relevant:

- In section 3.a), Mondays and Thursdays have less probability of using the oven as well as the number of openings above 40;
- In section 3.b), it is obvious that the case with more than 25 openings is not common and could be dropped.

In other words, we could discard observations with probability less than a specific threshold (which needs to be determined experimentally). Moreover, we can try predictions based on the existing data.

An attractive research direction is distributed data stream mining [12] in the home environment to extract patterns from data provided by multiple household appliances (stove, refrigerator, central heating), based on mobile agents technology [13] that can improve the usage of Big Data volumes.

ACKNLOWLEDGEMENT. The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement No609143 Project ProSEco.

REFERENCES

1. Merriam, Sharan B., *Qualitative Research and Case Study Applications in Education. Revised and Expanded from" Case Study Research in Education."*. Jossey-Bass Publishers, 350 Sansome St, San Francisco, CA 94104, 1998.

2. **Yilmaz, Özdoğan**. *Seismic data analysis. Vol. 1*. Tulsa: Society of Exploration Geophysicists, 2001.

3. Kenward, Robert. *Wildlife radio tagging: equipment, field techniques and data analysis*. London: Academic Press, 1987.

4. Jongman**, Robert HG, Cajo JF Ter Braak, and Onno FR Van Tongeren**, eds. *Data analysis in community and landscape ecology*. Cambridge university press, 1995.

5. **Darigny, Franck, Jacky Guerin, and Paul Muller**. *"Electronic control device for automatic cooking, including learning for home electric oven."* U.S. Patent No. 4,914,277. 3 Apr. 1990.

6. **Omura, Y., and M. Losco**. "*Electro-magnetic fields in the home environment (color TV, computer monitor, microwave oven, cellular phone, etc) as potential contributing factors for the induction of oncogen C-fos Ab1, oncogen C-fos Ab2, integrin alpha 5 beta 1 and development of cancer, as well as effects of microwave on amino acid composition of food and living human brain*."*Acupuncture & electro-therapeutics research* 18.1 pp. 33-73,1992.

7. **Li, Bojun, Piyanuch Hathaipontaluk, and Suhuai Luo**. "*Intelligent oven in smart home environment*." *Research Challenges in Computer Science, 2009. ICRCCS'09. International Conference on*. IEEE, 2009.

8. **Boulet, Micael, et al**. "*CFD modeling of heat transfer and flow field in a bakery pilot oven*." *Journal of food engineering* 97.3 pp. 393-402, 2010.

9. **Neamt Liviu**, "*Energy audit for domestic customers*", Energetica, year 62. no. 7, pp. 278-281, 2014.

10. Hedges, Larry V. "*Statistical Methodology in Meta-Analysis*", 1982.

11. **Hotelling, Harold**. "*Analysis of a complex of statistical variables into principal components*." *Journal of educational psychology* 24.6, 1993,.

12. **Costea, C**. and **Cola, C**. "*A Practical Model for Distributed Sensing and Analysis using RPi*", Proceedings of the International Conference of the Carpathian Euro-Region's Specialists in Industrial Systems, CEurSIS pp.180-183, 2014.

13. **Costea, C.** "*Applications and Trends in Mobile Cloud Computing*." Carpathian Journal of Electronic and Computer Engineering, vol 5, pp. 57-60., 2012.