



Cour des comptes





# Pisa Summer School 2019

## Digital Transformation of Audit

Examples from the Belgian Court of Audit

Denis GROFELS



# Use case : FOREM

- The « FOREM » is the Walloon office for vocational training and employment. It is a public service intended for all legal and natural persons involved in employment and training issues as well as for players in labour market in general.
- An audit was performed to control the administrative and financial situation of employees : status, financial and administrative seniority, financial scales, salary supplements, allowances and reimbursement of expenses, etc.

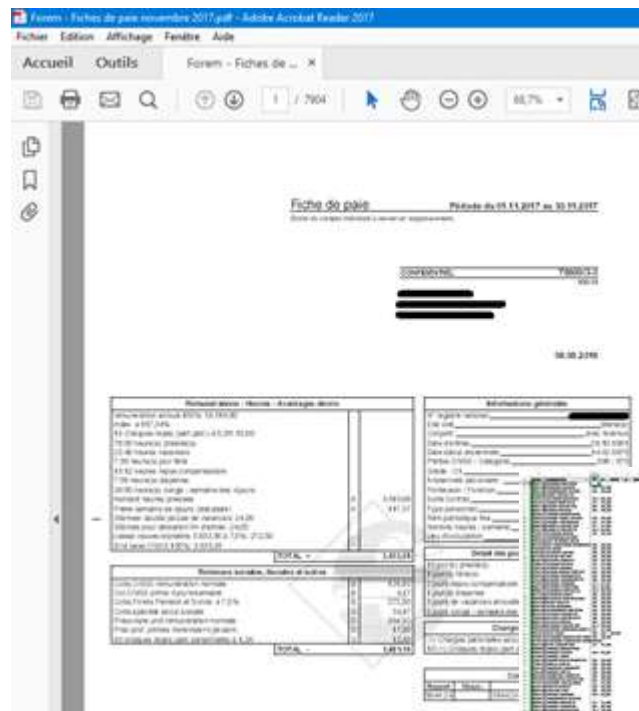


<b>Integration</b> Data acquired from various channels	<b>Processing</b> New information from available data	<b>Reporting</b> Production of thematic reports
<b>Data received directly in structured form</b> Example : Data about workforce, staff inflows and outflows Excel/Access	<b>Sampling of sub-populations defined from data analyses</b> Example : Staff hired with seniority greater than 10 years in round number of years.	Example : Staffing
<b>Manual structuring of documents</b> Example : Pay grades, salary grids Excel/Access	<b>Analyses of ratios computed from different sources</b> Example : Consistency of ratios between reimbursement of costs for home-work travel and home-work distance.	Example : Paygrades
<b>Data extraction from PDF documents</b> Example : Pay slips Java/PDFBox 7,904 PDF pages to a table of 208 columns x 4,697 lines	<b>Statistical analysis</b> Example : Frequency distribution of reimbursement of professional travel costs, recoded in a discrete number of classes.	Example : Salary supplements
<b>Automated data extraction from business application</b> Example : Individual sheets showing career histories AutoHotKey/VBA	<b>Anomaly detection</b> Example : Local Outlier Factor (LOF) used on 10 most important components of the payroll to identify atypical combinations. Python/Scikit-Learn	Example : Benefits in kind
<b>Web scrapping</b> Example : Distances from home to work Python/BeautifulSoup		Example : Contract staff → permanent staff

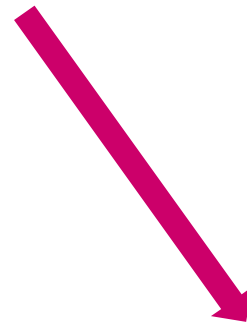
*Contexte* **Audit de la situation administrative et pécuniaire des agents du Forem**

*Cas d'utilisation* Extraction de texte

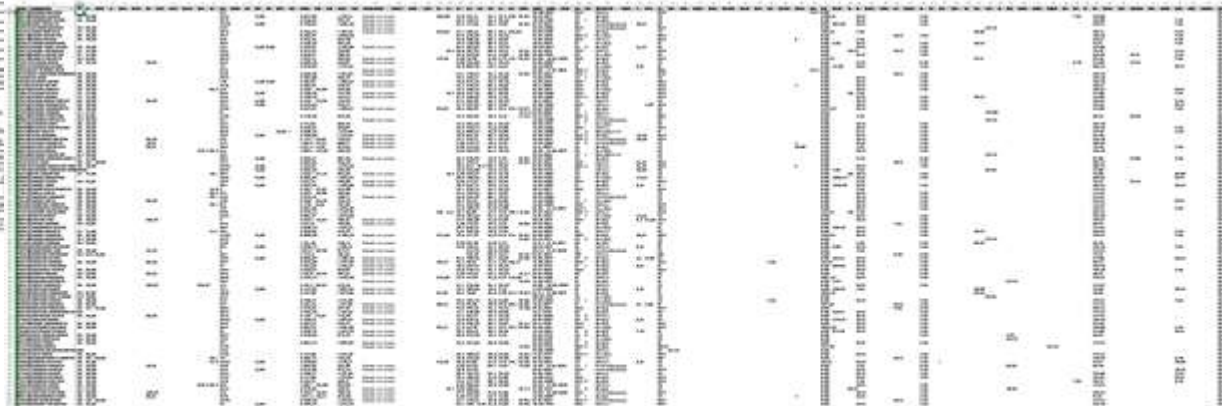
*Logiciels utilisés*  Java + PDFBox



7.904 pages en PDF



208 colonnes x 4.697 lignes en Excel





<i>Contexte</i>	<b>Audit de la situation administrative et pécuniaire des agents du Forem</b>
<i>Cas d'utilisation</i>	Webscraping
<i>Logiciels utilisés</i>	 Python + BeautifulSoup

**EXCEL**  
↓

**PDF**  
↓

**WEB**  
↓

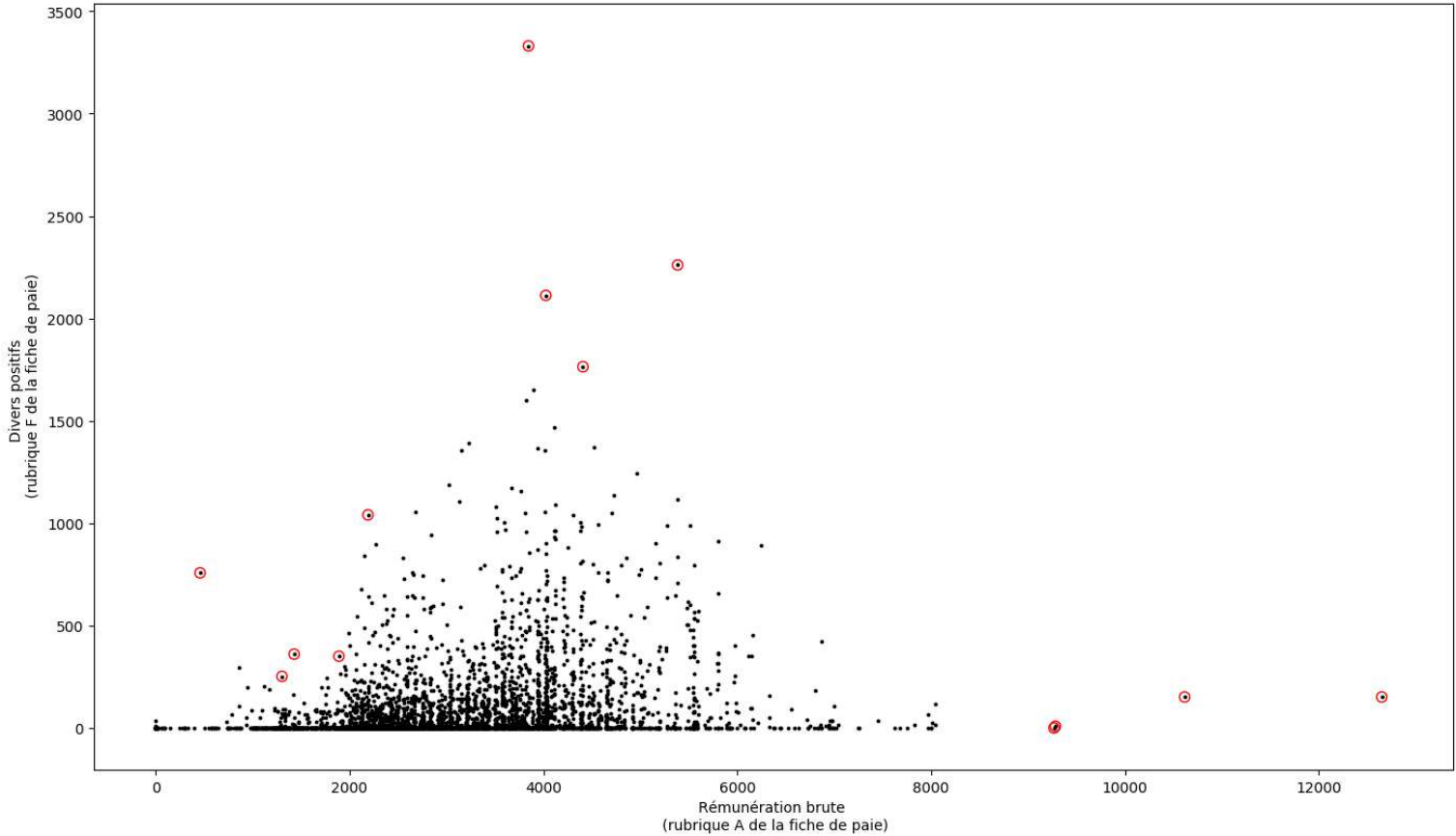
**RATIO**  
↓

NISS	Nom	Prénom	Commune de résidence	Lieu d'occupation	Distance	JPREST	FTRACOM	FTRACOM/KM
*****	*****	*****	Ottignies	Ottignies-L-L-Neuve	0,04	14	24,70	617,50
*****	*****	*****	Baulers	Nivelles	2,39	18	95,40	39,92
*****	*****	*****	Glain	Liège	2,28	17	73,50	32,24
*****	*****	*****	Glain	Liège	2,28	17	37,50	16,45
*****	*****	*****	Luingne	Mouscron	2,17	19	27,20	12,53
*****	*****	*****	Ougrée	Grâce-Hollogne	6,09	18	73,40	12,05
*****	*****	*****	Maisières	Mons	3,91	18	46,45	11,88
*****	*****	*****	Houtaing	Tournai	20,96	16	235,46	11,23
*****	*****	*****	Musson	Virton	12,50	15	137,60	11,01



<i>Contexte</i>	<b>Audit de la situation administrative et pécuniaire des agents du Forem</b>
<i>Cas d'utilisation</i>	Détection automatique de combinaisons de valeurs atypiques dans les fiches de paie
<i>Logiciels utilisés</i>	 Python + Scikit-learn

Fiches de paie de novembre 2017  
Détection de valeurs atypiques par la méthode "Local Outlier Factor" (LOF)





# Questions ?

**Denis Grofils**

*Auditeur adjoint*

Audits thématiques des communautés et des régions



Cour des comptes

Rue de la Régence 2, B-1000 Bruxelles

T +32 2 551 84 89

[GrofilsD@ccrek.be](mailto:GrofilsD@ccrek.be)

[www.courdescomptes.be](http://www.courdescomptes.be)