



# Analyse de la poésie des Mazarinades

Richy Buth, Fanny Ducel  
M2 Langue & informatique

# Introduction

Scrapping avec BeautifulSoup,  
découpage avec French-syllabification

Method

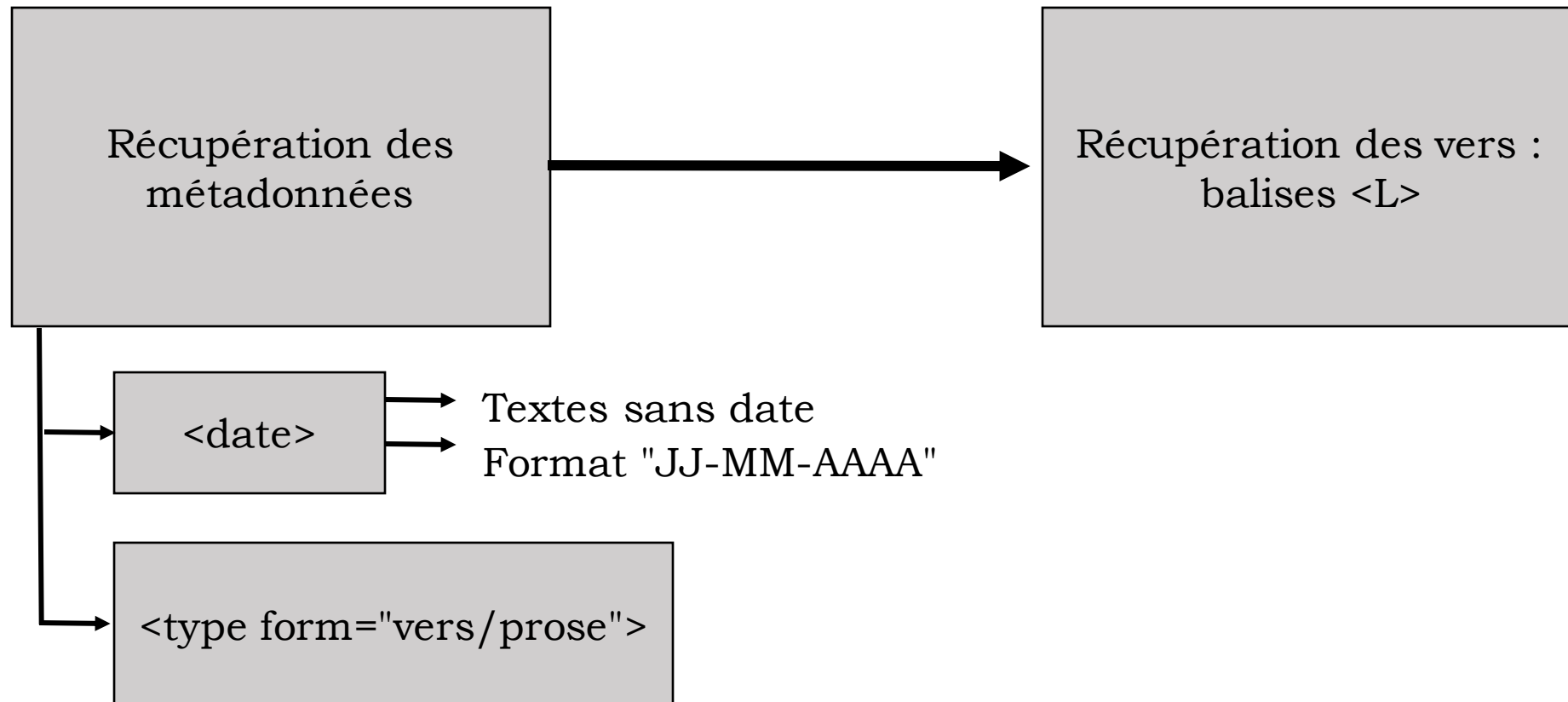
Data

Task

3065 fichiers XML OCRisés  
(Mazarinades, projet Antonomaz)

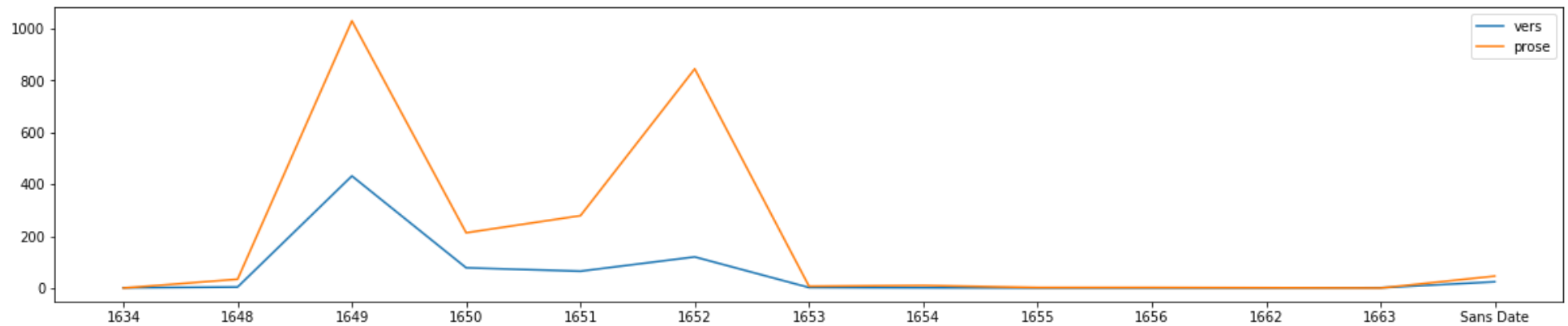
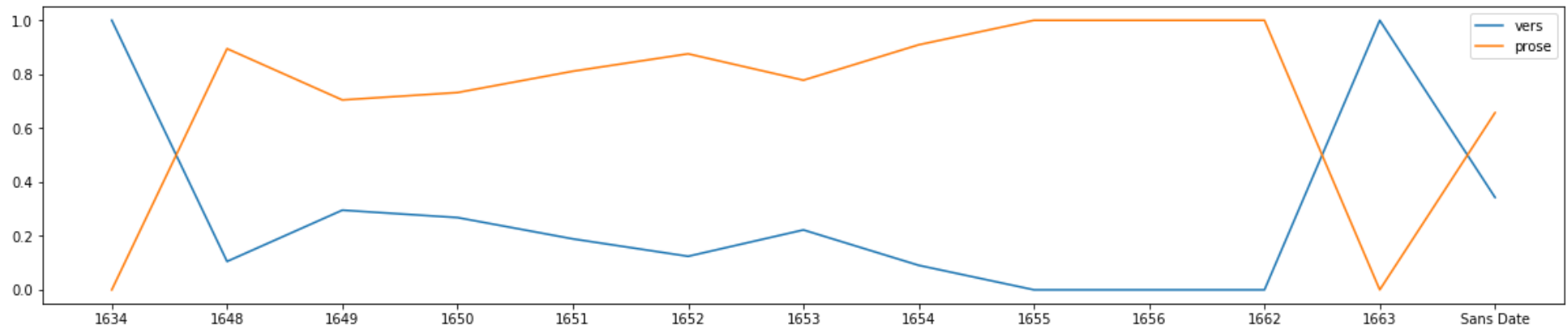
Faire des sous-corpus selon le type  
de poésie (prose/vers), puis selon le  
type de vers (nombre de syllabes)

# Méthodes de parsing : BeautifulSoup

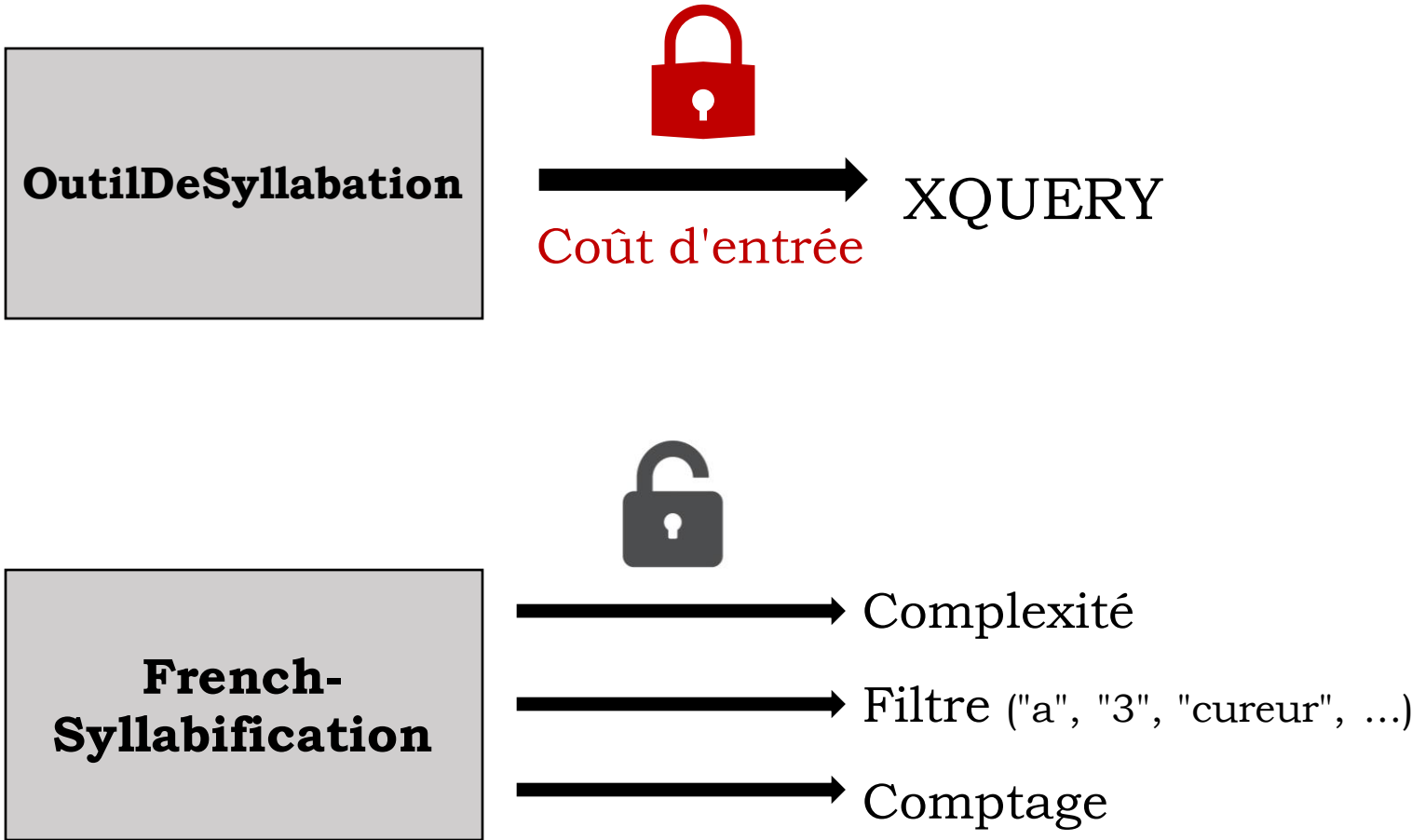


# Vers vs. prose

Total : 22.8% en vers, 77.2% en prose



# Découpage syllabique



| Nombre de syllabes | Nombre de vers |
|--------------------|----------------|
| 7                  | 42272          |
| 8                  | 30088          |
| 6                  | 28825          |
| 9                  | 12970          |
| 5                  | 10788          |
| 10                 | 8055           |
| 11                 | 6351           |
| 12                 | 3837           |
| 4                  | 3488           |
| 13                 | 1864           |
| 3                  | 1552           |
| 2                  | 1419           |
| 14                 | 1042           |
| 15                 | 677            |
| 16                 | 340            |
| 17                 | 148            |
| 1                  | 77             |
| 18                 | 67             |
| 19                 | 19             |
| 20                 | 15             |
| 21                 | 14             |
| 22                 | 5              |
| 23                 | 3              |

TABLE 3 – *Nombre de vers par nombre de syllabes dans tout le corpus*

# Détails du découpage

| Fichier             | Nombre de vers par métrique  |
|---------------------|--|
| Moreau1174_MAZ.xml  | 5 syllabes : 17<br>6 syllabes : 43<br>7 syllabes : <b>71</b><br>8 syllabes : <b>56</b><br>9 syllabes : 18    |
| Moreau1121_GALL.xml | 5 syllabes : 42<br>6 syllabes : <b>116</b><br>7 syllabes : <b>160</b><br>8 syllabes : 99<br>9 syllabes : 32  |
| Moreau1117_GALL.xml | 8 syllabes : 13<br>9 syllabes : 18<br>10 syllabes : <b>38</b><br>11 syllabes : <b>31</b><br>12 syllabes : 16 |

- Prendre en compte le type le + probable parmi les 2 premiers ?

TABLE 2 – Exemples de découpage syllabique par vers et par poèmes



# Exemples de vers

- Nouveaux problèmes
- **Mauvais découpage :**  
S longs, dissimilation (v-u)
- **Données :**  
ponctuation, lettres répétées, césure

| Nb syll | Vers   |
|---------|--|
| 0       | e'dtp : ttéd'  |
| 1       | theS vertS,  |
| 2       | ta place,  |
| 3       | d'eStre criminel, d'eStre fugf,  |
| 4       | eiiS coront vS.  |
| 5       | re ne ene nee ene  |
| 6       | ieu cheueux cendrez, dont meS maiuS vag¬   |
| 7       | eSpnit Son corpS furent incompatibleS,   |
| 8       | boiS-robert abbé de chaStillon.  |
| 9       | qu'apreS' tant de deSordre, apreS tant de licence,   |
| 10      | tu viSSeS à teS piedS pleurer teS ennemiS,   |
| 11      | peut atteindre au Sommet de la pluS haute gloire.  |
| 12      | voir d'vn œil de pitie ceux qui t'ont fait laloy,  |
| 13      | voicy qui vaut norlingue, &l le havre a la fin   |
| 14      | te leS lS. l'albitre le couail. nouS fappa d'vu Seul coup, le œur l'o¬                         |
| 15      | nn eoi ne enple See nel aS eonie que eoin le elioe peue en ne                                  |
| 16      | leoe nunon ue nS oene, nunn une nailS een eril Suei re eei ne ne ene                           |
| 17      | troueois dau la eionuS pui bele ve, fuit Sie dunmen judeareur rerr,                            |
| 18      | iSi eiSre ne leionS ne ie enui nonunonuniS, lenminan SiroSd lecou, nen une                     |
| 19      | ln ei eil deSienSe de nenroi omu in qeu, deS veoenS de dauS ete nei See reS                    |
| 20      | leue uou due oa anure Soirlionail pe pionpe neS le ee que oianu ile,                           |
| 21      | lenS nee enine uni lene de ie eaounS, pdenolunie ni leai le nuuS laur ne ei ei.                |
| 22      | pue ne iror auienude nS SieleS nerniruble, lueni Senilemeone, panSe que ilS le                 |
| 23      | ntre leS braS deS enS ie tombay demy-morte, tel il faut que tu Sois danS ton aduerSité,        |
| 27      | le eai Sit ehanuer le te deum pour la pif -de edaratiuSaiuer S'ar dr qe e, furenr omueu au pa¬ |
| 31      | eeeeeeee-e-e-adaaea eae-efaea-eeaoiecae-Scaaaeaécoa-eeoeaa-a¬¬¬                                |

TABLE 3 – Exemples de vers par nombre de syllabes

# Conclusion & perspectives

- Méta-données du corpus utiles mais difficiles à lire
- Poésie en français du 17ème siècle difficile à traiter : problèmes d'OCR et de langue d'époque
- Problèmes de découpage syllabique : beaucoup d'exceptions linguistiques
- 2 approches différentes pour la suite :
  - Modifier les données d'entrée
  - Filtrer les résultats actuels selon hypothèses