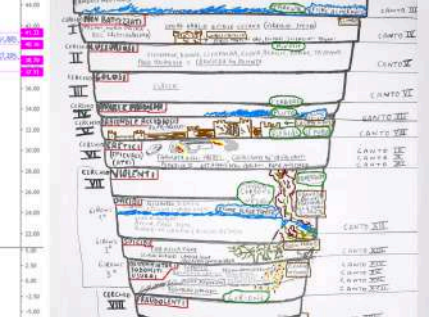
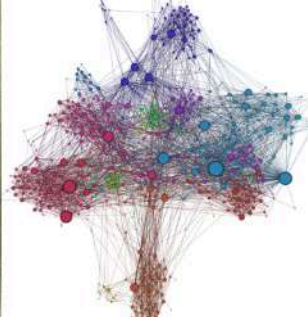


# For details, see <https://review.docs.microsoft.com/en-us/visualstudio/r/vs/sql-server>

```
# Test code
library(ROdbc)
channel <- odbcDriverConnect(dbConnection)
InputDataSet <- sqlQuery(channel, iconv(paste(readLines(
  'c:/proj/rproject1/rproject1/storedprocedure.query.sql',
  encoding = 'UTF-8', warn = FALSE), collapse = '\n'), from = 'UTF-8',
  to = 'ASCII', sub = ''))
odbcClose(channel)

OutputDataSet <- InputDataSet
```



# Présentation du cours (18 heures)

Giovanni Pietro Vitali – University College Cork

[giovannipetrovitali@gmail.com](mailto:giovannipetrovitali@gmail.com)

<https://github.com/digitalkoine>

<https://ucc-ie.academia.edu/GiovanniPietroVitali>



UCC

Coláiste na hOllscoile Corcaigh, Éire  
University College Cork, Ireland



# Objectifs

- Analyser un phénomène en extrayant les données nécessaires à une analyse distante
- Créer un dataset adapté à différents types d'analyses
- Connaître les logiciels, les applications et les langages les plus utilisés en *distant reading*
- Acquérir une pratique approfondie des analyses des réseaux, linguistiques et spatiales.
- Repenser un sujet dans le cadre d'une diffusion pour un public généraliste.

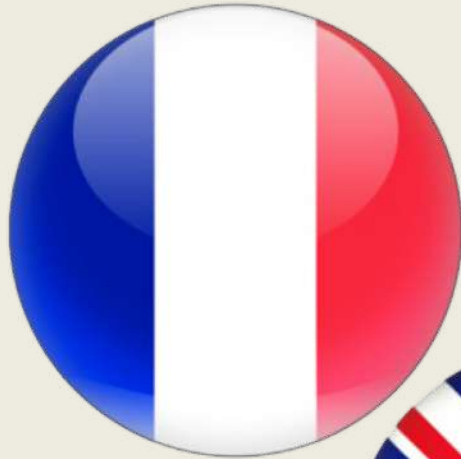


# Contenus abordés

- Stylométrie
- Traitement Automatique de la Langue
- Analyse du réseau
- Cartographie numérique

# Langage

☐ Français



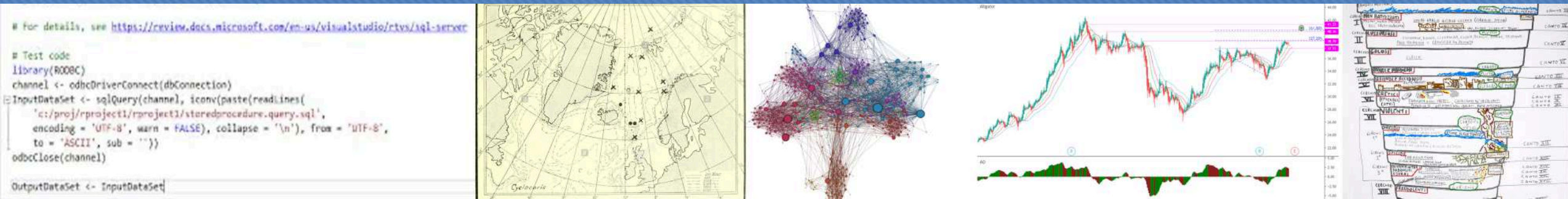
☐ English

# Logiciels, Applications et Langages

- R et Rstudio (Les packages seront installés lors du cours) <https://www.rstudio.com/products/rstudio/download/>
- Gephi <https://gephi.org/>
- Inkscape <https://inkscape.org>
- Compte Google <https://support.google.com/accounts/answer/27441?hl=fr>
- Compte GitHub education <https://help.github.com/en/articles/applying-for-a-student-developer-pack>
- Compte sur Carto à travers GitHub education <https://carto.com/blog/carto-is-part-of-the-github-student-pack/>
- QGIS (3.4.4 Madeira est la version conseillée) <https://qgis.org/en/site/forusers/download.html>
- TreeTagger <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>
- TXM <http://textometrie.ens-lyon.fr/?lang=fr>
- Microsoft Office et/ou OpenOffice <https://www.openoffice.org/download/>
- OpenRefine <http://openrefine.org/>



# Jour 1



## Introduction & Réseaux

11h-12h

Qu'est-ce que le distant reading ? Définition et analyse des exemples de recherche

12h-13h

Introduction à la sentiment analysis et aux outils de digital method initiative

14h-15h

Gephi et les réseaux

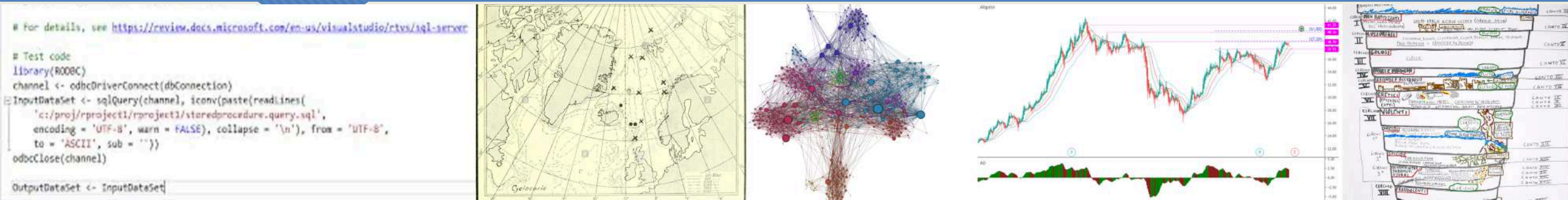
15h-16h

Un exemple de sentiment analysis extrait grâce aux outils de digital method initiative et élaboré avec Gephi

16h-17h

Exercice sur Gephi

# Jour 2



## Travailler avec les textes

11h-12h

Stylométrie et Textométrie, usages et méthodologies

12h-13h

Stylométrie avec R, les premiers pas

14h-15h

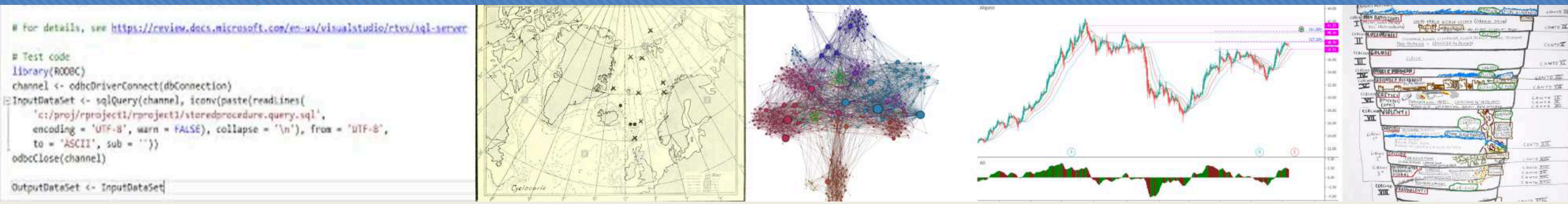
Stylométrie avec R, les premiers pas

15h-16h

Customiser les représentations stylistiques par l'usage de Gephi et 3dNetwork



# Jour 3



## Du texte à la carte géographique

11h-12h

TXM & TreeTagger : une introduction aux lemmatiseurs

12h-13h

Analyses linguistiques et sémantiques par le lemmatiseur

14h-15h

Extraire les toponymes des textes

15h-16h

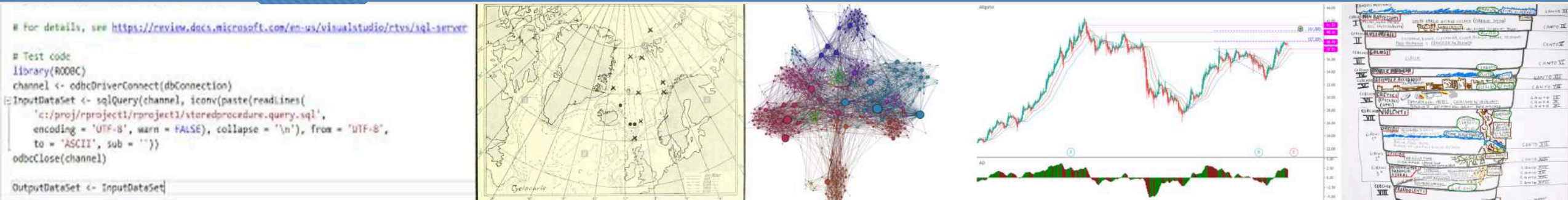
Géoréférencer, un panorama des options

16h-17h

Spatialiser les données : ma première carte en ligne avec Carto



# Jour 4



## Cartographie numérique

11h-12h

Créer une carte avec R : préparation des données

12h-13h

Créer une carte avec R : les points

14h-15h

Créer une carte avec R : les polygones

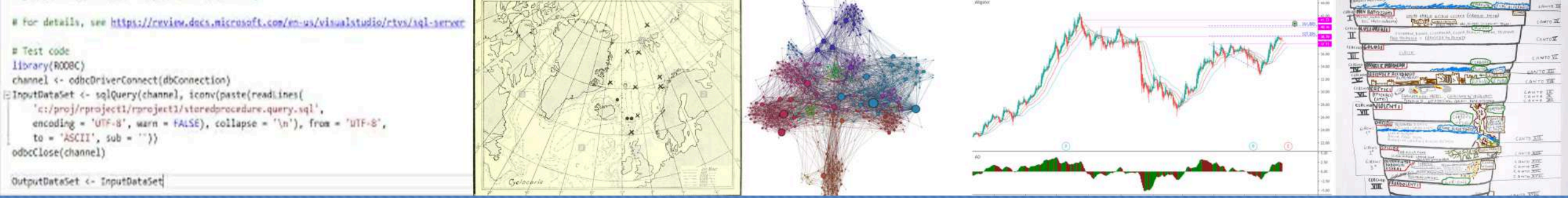
15h-16h

Créer une carte avec R : le graphique

# Évaluation

- L'évaluation de ce cours se fera sur la base de l'élaboration d'un dossier de recherche portant sur l'application d'au moins deux des disciplines présentées tout au long du cours.
- Les étudiants devront préparer au moins une analyse en *distant reading* d'un sujet de leur choix (après concertation avec l'enseignant) et écrire un dossier de six pages minimum, dont trois seront dédiées aux aspects techniques des outils numériques utilisés, et trois aux conclusions interprétatives permises par l'utilisation de ces outils. L'usage combiné des divers outils comme, par exemple, l'analyse des réseaux avec le traitement automatique de la langue, sera fortement apprécié.
- Le dossier et les analyses numériques seront présentés dans le cadre d'un entretien individuel de 20 minutes (10 minutes d'exposé et 10 minutes de questions).





# Présentation

# \_end

Giovanni Pietro Vitali – University College Cork

[giovannipetrovitali@gmail.com](mailto:giovannipetrovitali@gmail.com)

<https://github.com/digitalkoine>

<https://ucc-ie.academia.edu/GiovanniPietroVitali>



UCC

Coláiste na hOllscoile Corcaigh, Éire  
University College Cork, Ireland

