

Année scolaire 2021-2022

Master « Mathématiques Appliquées, Statistique »

parcours-type

Evaluation et décision publiques

délivré par



Campus de Ker Lann
Rue Blaise Pascal
BP 37203
35172 Bruz cedex
Tél. : 02 99 05 32 32
www.ensai.fr

en cohabilitation avec



Faculté des sciences économiques
Campus centre
7 place Hoche
35065 Rennes cedex
Tél. : 02 23 23 35 45
www.eco.univ-rennes1.fr

Contenu

PRESENTATION GENERALE	5
TRONC COMMUN	9
CONCEPTION D'ENQUETES	10
ECONOMETRIE AVANCEE	12
ANALYSE ET TRAITEMENT DES SÉRIES TEMPORELLES.....	14
ECHANTILLONNAGE AVANCE	15
EMOS	17
DEVELOPPEMENT D'APPLICATIONS	44
PROFESSIONNALISATION	18
ANGLAIS	19
<i>ENGLISH</i>	19
DEMOGRAPHIE.....	21
ANALYSE TEXTUELLE	22
ENTREPRISE	23
THEORIE DES INDICES DE PRIX.....	24
ÉVALUATION D'IMPACT DES POLITIQUES PUBLIQUES : DES EXEMPLES PRATIQUES	26
REGRESSION PENALISEE	ERREUR ! SIGNET NON DEFINI.
STAGE	28
MAJEURE « METHODOLOGIE DE LA STATISTIQUE PUBLIQUE ».....	29
SERIES TEMPORELLES AVANCEES.....	30
MODELES DE PREVISION MULTIVARIES	31
ENQUÊTES LONGITUDINALES.....	32
ESTIMATION SUR PETITS DOMAINES ET COMPLEMENTS SUR LES REDRESSEMENTS	33
CLASSIFICATION NON SUPERVISEE ET REDUCTION DE DIMENSIONS.....	34
MULTICOLINEARITE ET REGRESSION PLS	36
APPRENTISSAGE, AGREGATION DE MODELES ET SVM.....	38
ANALYSE SPATIALE	40
ANALYSE DE DONNEES D'ENQUETES	42
MAJEURE « STATISTIQUE ET TRAITEMENT DES DONNEES ».....	43
RESEAUX, SYSTEMES D'EXPLOITATION, SECURITE DE L'INFORMATION	46
ECHANGES SUR LE WEB.....	48
MASSES DE DONNEES DISTRIBUEES	49
PARCOURS « ETUDES STATISTIQUES ».....	51
EVALUATION DES POLITIQUES PUBLIQUES	52
ETUDES DE CAS DE POLITIQUE PUBLIQUE	53
ANALYSE DE LA CONJONCTURE.....	55

Présentation Générale

Objectifs

Les acteurs publics disposent de bases de données locales, nationales voire internationales, dont la multiplication a été favorisée par le développement des systèmes d'information ; ils ont donc besoin de statisticiens qualifiés au plus haut niveau pour les concevoir, les analyser et les exploiter. Le but de ce parcours-type du master « mathématiques appliquées -statistiques » est précisément de répondre à cette demande croissante en expertise statistique pour l'aide à la décision publique.

Ce master bénéficie du label European master in official statistics (EMOS) délivré par Eurostat.

Quels métiers ? Quels secteurs d'activité ?

D'une manière générale, la formation reçue prépare les diplômés aux métiers scientifiques et techniques centrés sur l'ingénierie statistique, les systèmes d'information économique et sociale et les métiers informatiques liés à la statistique publique.

Elle répond tout d'abord aux besoins des services publics, administrations nationales ou collectivités locales. Les principaux ministères (économie, agriculture, développement durable, santé et action sociale, etc.), des établissements publics (notamment les caisses de sécurité sociale et les organismes gestionnaires de prestations), les organismes consulaires et les agences dépendant des collectivités locales sont susceptibles de recourir aux compétences des titulaires de ce master. Ils pourront aussi trouver des opportunités dans des organismes européens (Eurostat) et internationaux (OCDE, FMI, Banque mondiale).

Par ailleurs, les décideurs publics, notamment au niveau local, peuvent s'appuyer sur des cabinets de conseil spécialisés dans la décision publique : ces cabinets peuvent ainsi également offrir des débouchés aux titulaires du master de statistique publique.

Enfin, les attachés statisticiens de l'Insee peuvent compléter leur formation initiale par ce master.

Conditions d'admission

La formation est ouverte aux élèves fonctionnaires de l'Ensaï admis à la formation complémentaire diplômante, ainsi qu'aux étudiants ayant suivi la première année du master ; la formation est également ouverte aux diplômés de niveau équivalent. La sélection se fait par examen du dossier et un entretien. La formation est également ouverte en formation continue.

Calendrier de candidature : d'avril à fin mai.

Effectif : 25.

Pour tout renseignement sur l'admission : admission-statpublique@ensai.fr

Responsables de la formation

A l'université Rennes 1 : Isabelle Cadoret – isabelle.cadoret-david@univ-rennes1.fr

A l'Ensai : Brigitte Gelein – brigitte.gelein@ensai.fr

Formation continue diplômante des attachés (FCDA) et scolarité du master : Didier Le Bozec – didier.le-bozec@ensai.fr

Organisation des études

Beaucoup d'enseignements sont associés à la réalisation de travaux de groupes et de travaux personnels sous forme de dossiers à rendre.

La professionnalisation est au cœur du master. La majorité des enseignements est dispensée par des acteurs du système statistique public.

L'anglais est obligatoire en formation initiale.

Majeure Méthodologie de la statistique publique (MSP)	ECTS	COURS	TD/TP	TOTAL
SEMESTRE 1	20	114	66	180
UE 1 – Conception d'enquêtes	4	18	18	36
UE 2 – Econométrie avancée	4	24	12	36
UE 3 – 3.1 Analyse et traitement des séries temporelles – 3.2 Echantillonnage avancé	4	24	12	36
UE 4 – EMOS	4	24	12	36
UE 5 – Cours au choix parmi M1 à M6	4	24	12	36
SEMESTRE 2	40	48/64	84/72	132/136
UE 6 – Cours au choix parmi M1 à M6	4	24	12	36
UE 7 - Cours au choix parmi M1 à M6 et E1 à E3	4	24/40	12/0	36/40
UE 8 – Professionnalisation	2		60	60
UE 9 – Stage	30			
TOTAL	60	162/178	150/138	312/316

Majeure Statistique et traitement des données (STD)	ECTS	COURS	TD/TP	TOTAL
SEMESTRE 1	20	91	105	196
UE 1 – Conception d'enquêtes	4	18	18	36
UE 2 – Econométrie avancée	4	24	12	36
UE 3 – 3.1 Analyse et traitement des séries temporelles – 3.2 Echantillonnage avancé	4	24	12	36
UE 4bis – Développement d'applications S. 4.1 – Génie logiciel S. 4.2 – Développements d'applications web	4	10	38	48
UE 5 – Réseaux, systèmes d'exploitation, sécurité de l'information S. 5.1 – Réseaux S. 5.2 – Architecture et systèmes d'exploitation S. 5.3 – Cybersécurité	4	15	25	40
SEMESTRE 2	40	20	136	156
UE 6 – Echanges sur le web S. 6.1 – Les éléments de base d'une application web S. 6.2 – Communication entre applications web	4	10	38	48
UE 7 – Masses de données distribuées S. 7.1 – Développement d'applications distribuées S. 7.2 – Bases des données en grande dimension	4	10	38	48
UE 8 – Professionnalisation	2		60	60
UE 9 – Stage	30			
TOTAL	60	111	241	352

Majeure Etudes Statistiques (ES)	ECTS	COURS	TD/TP	TOTAL
SEMESTRE 1	20	130	54	184
UE 1 – Conception d’enquêtes	4	18	18	36
UE 2 – Econométrie avancée	4	24	12	36
UE 3 – 3.1 Analyse et traitement des séries temporelles 3.2 Echantillonnage avancé / Méthodes statistiques avancées(*)	4	24	12	36
UE 4 – EMOS	4	24	12	36
UE 5 – Evaluation des politiques publiques (E1)	4	40		40
SEMESTRE 2	20	76/80	60	136/140
UE 6 – Cours au choix parmi E2 et E3 (**)	4	40		40
UE 7 – Cours au choix parmi M1 à M6 et E2 à E3 (**)	4	36/40		36/40
UE 8 – Professionnalisation	2		60	60
UE 9 – Stage	30			
TOTAL	60	206/210	114	320/324

Pour les étudiants inscrits à l’Université de Rennes 1 :

(*) l’enseignement suivi au titre de l’UE 3.2 est « Méthodes statistiques avancées ».

(**) les enseignements suivis au titre des UE 6 et UE 7 sont respectivement E2 et E3.

<i>Enseignements de choix</i>
M1 : M11 Séries Temporelles Avancées + M12 Modèles de Prévion Multivariés
M2 : M21 Enquêtes longitudinales + M22 Estimation sur petits domaines et compléments sur les redressements
M3 : M31 Classification non supervisée et réduction de dimensions + M32 Multicolinéarité et régression PLS
M4 : M41 Apprentissage, agrégation de modèles et SVM + M42 Analyse Spatiale
M5 : M51 Analyse de données d’enquêtes + M42 Analyse Spatiale
M6 : M31 Classification non supervisée et réduction de dimensions + M41 Apprentissage, agrégation de modèles et SVM
E1 : Evaluation des politiques publiques
E2 : Etude de cas de politique publique : le marché du travail - Etude de cas de politique publique : l’environnement
E3 : Analyse macro-économique de la conjoncture - Analyse du marché des matières premières

Contrôle des connaissances

L’évaluation des enseignements est précisée dans chaque présentation d’enseignement.

Règlement intérieur

Le règlement intérieur de la scolarité est celui de l’Ensaï.

Tronc commun

Conception d'enquêtes

Survey methodology

Cours : 18h • TD/TP: 18h

Enseignant : Amandine Schreiber (Insee), Thomas Amossé (Cnam), Stéphane Legleye (Insee)

Correspondant : Brigitte Gelein

Objectif : Cette formation propose d'analyser le travail de conception d'enquête, notamment à travers différents exemples d'enquêtes. La phase de rédaction du questionnaire est détaillée. Une participation active des étudiants est sollicitée.

Contenu : les différents points ci-dessous seront abordés en cours ou en TD.

La conception d'enquête, entre statistique et sociologie

Un travail sociologique : identifier une problématique, un champ, une population. Pourquoi quantifier, comment mesurer : pour quels objectifs, avec quels indicateurs, quelles catégories (définition du ménage, des catégories sociales, types d'entreprises, etc.)

Le travail de statisticien d'enquête, de l'échantillonnage au traitement statistique
Spécificités et similitudes des enquêtes ménages et des enquêtes entreprises

Un travail collectif impliquant de nombreux métiers et partenaires

L'équipe de conception

Des acteurs clés : les enquêteurs

Comitologie : comité scientifique, comité de pilotage et/ou de suivi, comité éthique, etc.

Les acteurs institutionnels : Cnis, comité du Label, Cnil

Les partenaires extérieurs, les sous-traitants éventuels

La gestion de la collecte

Le questionnaire

Les différents modes de collecte

Les principaux risques de biais et erreurs de mesure

Règles de rédaction des questions

Le rôle des tests

Figures d'enquêtes particulières

Enquêtes comparatives internationales (GGS, Pisa...)

Enquêtes complexes : enquêtes auprès des sans domicile, enquêtes couplées ménages/entreprises, enquêtes sur la santé et le handicap...

Panels et enquêtes longitudinales

Enquêtes présentées en TD

Travaux dirigés

Travail d'écriture de petits modules de questions

Présentation d'articles par groupe de 2 ou 3 étudiants et discussion collective

Présentation détaillée d'enquêtes par groupe de 2 ou 3 étudiants et discussion collective

Contrôle des connaissances

L'évaluation reposera sur trois dimensions :

- participation active au cours
- présentation d'un article
- présentation d'une enquête

La note finale est la moyenne des notes obtenues pour les présentations de l'article et de l'enquête pondérée par la participation active au cours.

Références bibliographiques

- *Méthodes et pratiques d'enquêtes*, 2010, Statistique Canada.
- Insee (2009 a et b), Dossier spécial : concevoir une enquête auprès des ménages, *Courrier des statistiques*, n° 126 et 128, jan./fév. et sept./déc
- François de Singly, 2002, *L'enquête et ses méthodes : le questionnaire*, collection sociologie, Éditions Nathan.

Langue d'enseignement

Français

Econométrie Avancée

Cours : 24 TD : 12

Enseignant : Isabelle Cadoret, Véronique Thelen (Université Rennes 1), François Coquet (Ensaï)

Correspondant : Isabelle Cadoret (Université Rennes 1)

Objectif de l'UE

L'objectif de l'UE est de former les étudiants à l'utilisation de techniques économétriques avancées en s'appuyant sur les connaissances de base du modèle de régression. L'accent est mis sur l'économétrie l'économétrie spatiale et l'économétrie des panels.

Le cours d'économétrie spatiale s'appuie sur les connaissances de base du modèle de régression en les appliquant à des données localisées. Une des caractéristiques de ces données géo-spatiales est de créer une dépendance spatiale entre les observations, une interaction spatiale, qui doit être prise en compte dans l'analyse économétrique. En effet, l'action d'un acteur (un pays, une région, une entreprise, un acheteur dans l'immobilier, etc.) n'est pas indépendante de celle de ses voisins localisés à une plus ou moins grande distance, et celle-ci est source d'externalités qui dépendent du degré de proximité entre les observations.

La prise en compte des effets externes, des effets de réseaux ou encore des interactions stratégiques entre agents économiques est importante dans différents domaines : économie urbaine, analyse de la croissance économique, économie publique, économie de l'innovation, économie internationale...

Le cours d'économétrie des panels s'appuie également sur le modèle de régression avec des données individuelles et temporelles. Les données de panel permettent d'analyser la dynamique de comportement d'individus hétérogènes. La prise en compte de l'hétérogénéité est fondamentale dans la modélisation et l'objectif est de présenter les différents modèles.

A l'issue de l'UE l'étudiant(e) doit être capable de :

- Vérifier l'importance de la dimension spatiale dans une problématique
- Tester l'hétérogénéité des comportements des individus
- Choisir le modèle d'économétrie spatiale adapté et appliquer les méthodes d'estimation pour l'étude de cas concrets ;
- Utiliser des logiciels appropriés
- Interpréter les résultats et en dégager les implications
- Communiquer sur ces travaux à l'écrit (production de notes d'études et de recherche à un public spécialisé ou non) comme à l'oral

Partie 1 Econométrie spatiale

Contenu de la matière

Chapitre 1 : structuration des liens spatiaux et matrice de pondérations spatiales

Chapitre 2 : Statistiques descriptives et tests

Chapitre 3 : Modèles d'économétrie spatiale dans le cas linéaire

Chapitre 4 : Autres modèles d'économétrie spatiale

Partie 2 Econométrie des panels

Contenu de la matière

Chapitre 1: Analyse de la variance

Chapitre 2: Modèle à coefficients fixes : méthodes d'estimation et tests statistiques

Chapitre 3: Modèle à effets individuels aléatoires : méthodes d'estimation et tests statistiques

Chapitre 4 : Autre modèles d'économétrie des panels

Langue d'enseignement

Français

Pré-requis

Le modèle de régression linéaire simple et les modèles de séries temporelles. Méthode d'estimation MCO, MCG, GMM, par maximum de vraisemblance et tests associés.

Contrôle des connaissances

Un TP individuel et un projet en binôme

Références bibliographiques

- Anselin L. (2001) Spatial econometrics, dans Baltagi B. (ed.), *Companion to Econometrics*, Basil Blackwell, Oxford.
- Dubé J. et D. Legros (2014) Econométrie spatiale appliquée des microdonnées, ISTE ed.
- Le Gallo J. (2004) Hétérogénéité spatiale, principes et méthodes, *Economie et Prévision*, vol. 162, pp. 151-172.
- Patrick Sevestre, « Econométrie des données de panel » Dunod
- Alain Pirotte , « Econométrie des données de panel » Economica
- W. Greene, "Econometric Analysis", Prentice Hall.
- Cameron and P. Trivedi, "Microeconometrics : methods and applications", Cambridge University Press

Analyse et traitement des séries temporelles

Time Series

Cours : 12h • TD : 6h

Enseignants : Pierre Leblanc (Insee)

Correspondant : Brigitte Gelein

Objectif de la matière

Comprendre les enjeux particuliers associés à l'étude des séries temporelles ; maîtriser les outils essentiels d'analyse et de traitements pour la production ou l'exploitation de données temporelles.

Contenu de la matière

1. Rappels de séries temporelles ; outils d'études et d'analyse de la structure d'une série temporelle, filtre
2. Modèles univariés ARMA
3. Introduction à la non stationnarité, modèles ARIMA et SARIMA
4. Exemples de non linéarités et traitements associés
5. Traitement de la saisonnalité et des effets de jours ouvrables
6. Introduction aux méthodes de prévision

Pré-requis

Connaissance de base sur les séries temporelles et les modèles univariés (il est essentiel d'avoir suivi un premier cours sur le sujet).

Bonne compréhension de la régression linéaire.

Contrôle des connaissances

Un examen écrit

Références bibliographiques

D. Peña, G. C. Tiao, Ruey S. Tsay (2001), *A Course in Time Series Analysis*

J. D. Hamilton (1994), *Time Series Analysis*.

D. Ladiray, B. Quenneville (1999), *Comprendre la méthode X11*

V. Gomez, A. Maravall (2001), *Automatic modelling methods for univariate series*

V. Gomez, A. Maravall (2001), *Seasonal adjustment and signal extraction in Economic time series*

B. Salanié (1999), *Guide pratique des séries non-stationnaires*

Langue d'enseignement

Français

Echantillonnage avancé

Cours : 12h • TD : 6h

Enseignants : David Haziza (Université de Montréal), Guillaume Chauvet (Ensay)

Correspondant : Brigitte Gelein

Objectif de la matière

Après quelques rappels sur les principes généraux d'inférence et d'estimation dans le cadre d'une population finie, les méthodes d'échantillonnage à probabilités inégales seront présentées avec des exemples concrets d'utilisation.

Dans une seconde partie, la notion d'échantillonnage équilibré sera introduite. On étudiera en détail la méthode du Cube permettant de sélectionner des échantillons équilibrés, ainsi que l'estimation de variance pour les estimateurs associés. Là encore, on illustrera l'utilisation de cette méthode en se basant sur des exemples d'enquêtes.

Dans une dernière partie, on étudiera comment les méthodes d'échantillonnage avancé peuvent être utilisées au stade de l'imputation afin de réduire la variance des estimateurs imputés. Après quelques rappels sur l'imputation de la non-réponse partielle, on étudiera l'utilisation de méthodes d'échantillonnage équilibré permettant d'améliorer la précision des méthodes d'imputation aléatoires classiques.

Les différentes techniques présentées dans le cours seront travaillées dans le cadre de TP.

Contenu de la matière

Partie 1 : Rappels généraux sur l'échantillonnage

Partie 2 : Méthodes d'échantillonnage à probabilités inégales

- Tirage de Poisson
- Tirage à probabilités inégales de taille fixe
- Application au tirage autopondéré dans les enquêtes ménage

Partie 3 : Echantillonnage équilibré

- Principe
- Mise en œuvre
- Estimation de variance
- Application de l'échantillonnage équilibré

Partie 4 : Imputation équilibrée

- Rappel sur les méthodes d'imputation
- Imputation aléatoire équilibrée
- Mise en œuvre

Pré-requis

Une bonne maîtrise des techniques de sondages telles qu'elles sont présentées dans le cours de 2^{ème} année : Théorie des sondages, Données manquantes et Méthodologie d'Enquête.

Contrôle des connaissances

Le contrôle des connaissances se fait sous la forme d'un examen écrit ou d'un TP noté.

Références bibliographiques

- ARDILLY P. (2006) , *Les Techniques de Sondage*, Paris, Technip.
- SÄRNDAL, C-E., SWENSSON, B, WRETMAN, J. (1992), *Model Assisted Survey Sampling*, Springer-Verlag.
- TILLÉ Y. (2006), *Sampling Algorithms*, Wiley.

Langue d'enseignement

Français

EMOS

Cours : 24h • TP : 12h

Enseignant : Nathalie Shlomo (Université de Manchester), Kim Voldby Pedersen (Statistique Danemark)

Correspondant : Brigitte Gelein

Objectif de la matière

Contenu de la matière

Aims of Module

The two day course will introduce basic and advanced concepts of statistical disclosure control, privacy and confidentiality. The topics covered include the motivation of statistical disclosure control in terms of disclosure risk scenarios and types of disclosure risk; measuring disclosure risk for traditional outputs: microdata and tabular data; common methods of statistical disclosure control applications; the impact of statistical disclosure control methods on utility. In addition, we introduce differential privacy, a mathematical rigorous definition of a perturbation mechanism that was developed by computer scientists, which provides formal and quantifiable guarantees of confidentiality. Differential privacy is currently being explored by statisticians working within statistical agencies as we move towards more advanced and flexible modes of data dissemination.

Lecture Programme

- Introduction and motivation of statistical disclosure control for statistical outputs: disclosure risk scenarios, types of disclosure risk
- Measuring and quantifying disclosure risk for tabular data, microdata and other forms of disseminated outputs
- Disclosure control methods for statistical outputs
- Measuring and quantifying the impact and the effect of disclosure control methods on the quality of the data
- The differential privacy standard
- New forms of data dissemination and future challenges in statistical disclosure control

Pré-requis

Niveau d'anglais suffisant pour comprendre le cours et prendre la parole

Contrôle des connaissances

A définir

Références bibliographiques

Langue d'enseignement

Anglais

Professionnalisation

TD/TP : 60h

Enseignants : Divers intervenants

Correspondants : Divers

Anglais (24 h)

Matière obligatoire pour les élèves en formation initiale. Dispense pour les élèves en formation continue détenteurs du niveau B2 du ToEIC.

Cours d'ouverture (36 h)

« Démographie » (24 h) – correspondant : Brigitte Gelein

« Analyse textuelle » (12 h) – correspondant : Brigitte Gelein

« Entreprise » (12 h) – correspondant : Brigitte Gelein

« Indice des prix » (12 h) – correspondant : Brigitte Gelein

« Évaluation d'impact des Politiques Publiques : des exemples pratiques » (12 h) – correspondant : Brigitte Gelein

« Techniques de l'exposé » (12h) – correspondant : Brigitte Gelein

Au choix : démographie + l'une des 6 autres matières ou bien 3 matières (hors démographie) sur les 6. Les analystes sont dispensés de ces cours.

Anglais

English

Cours : 24h

Enseignants : Divers intervenants

Correspondant : Todd Donahue

Objectif de la matière

Cette formation obligatoire cible le développement et l'approfondissement des compétences en expression écrite et production orale nécessaires pour affronter le monde professionnel et scientifique en anglais. Elle cible également le développement et le perfectionnement des compétences grammaticales et lexicales.

Contenu de la matière

Selon les effectifs les cours seront organisés par groupe de niveau.

Les apprenants travailleront l'écriture des textes en langage soutenu (par ex. abstract, executive summary, synopsis) afin de résumer et/ou synthétiser les documents et rapports (articles de presse, rapports scientifiques issus des institutions et organismes tel que l'INSEE, Eurostat, OCDE, etc.).

Ce cours obligera les élèves de travailler l'expression orale et favorisera le développement de l'aisance (discussions/débats, présentation professionnelles...).

Pré-requis

Le niveau B1+ en anglais est conseillé.

Contrôle des connaissances

La note finale se calcule ainsi :

- Participation active en classe : 30 %
- Production écrit 1 (*abstract*) : 10 %
- Production écrit 2 (*executive summary/synopsis*) : 30 %
- Présentation orale : 30 %

Références bibliographiques

- <https://ec.europa.eu/eurostat/home>
- <https://www.oecd.org/france/>
- <https://www.imf.org/external/index.htm>
- <http://www.worldbank.org/>
- <https://www.insee.fr/en>
- Autres références sur l'écriture scientifique et professionnel en anglais seront données en cours

Langue d'enseignement

Anglais

Compétences Ciblées : B2 à C1 du [CECR](#)

- **Compétence de communication :**

- Ecouter/Lire

- Pouvoir comprendre les discours assez longs, les arguments complexe d'un sujet familier
 - Pouvoir comprendre des rapports scientifique/technique et des articles dans lesquelles les auteurs adoptent un certain point de vue

- Ecrire/Parler

- Pouvoir s'exprimer de façon claire et détaillée dans une variété de contextes
 - Pouvoir résumer, synthétiser, analyser et transmettre les informations complexes
 - Pouvoir paraphraser, de manière simple, un texte écrit
 - Pouvoir écrire des textes clairs, bien structurés et détaillés
 - Pouvoir participer activement à une conversation
 - Pouvoir présenter, expliquer et défendre ses opinions ou les opinions des autres
 - Pouvoir communiquer à l'oral avec un degré de spontanéité et d'aisance

- **Compétences linguistiques**

- Grammaire/syntaxe

- Pouvoir s'exprimer sans faire des erreurs systématiques
 - Pouvoir maîtriser la syntaxe de la phrase simple et utiliser des structures/phrases complexes
 - Pouvoir démontrer une maîtrise suffisante de la grammaire pour :
 - permettre une lecture fluide
 - ne pas faire d'erreurs conduisant à des malentendus

- Lexique

- Pouvoir utiliser un lexique pertinent et varié (éviter la répétition)
 - Pouvoir éviter des erreurs de lexique conduisant à des malentendus (par ex. les faux-amis)

- Phonologique :

- Pouvoir s'exprimer avec une prononciation et une intonation suffisamment claire pour être compris

Démographie

Cours : 24h

Enseignant : Laurent Di Carlo

Correspondant : Brigitte Gelein

Objectif de la matière

Cet enseignement vise à présenter les grandes tendances démographiques, à en comprendre les causes, conséquences et principaux outils de mesure et à savoir utiliser la démographie pour mieux comprendre un territoire. L'enseignement comporte une part importante de pratique, avec la construction en fil rouge d'un bilan démographique complet et sa présentation orale en s'appuyant sur un diaporama.

Contenu de la matière

Les tendances démographiques récentes

Urbanisation

Dividende démographique

Les migrations internationales

Définition, mesure

Grands axes de migration

Motifs et conséquences des migrations

Réfugiés et déplacés

Les projections de population

Eléments pour la construction d'un bilan démographique

Les indicateurs de croissance démo (SM, SN...)

Les indicateurs de fécondité

Le vieillissement de la population

La mortalité

Indicateurs et table de mortalité

Espérance de santé

Soutenance cas pratique

Et en fil rouge : construction progressive d'un bilan démographique avec analyse des indicateurs et présentation des résultats

Pré-requis

Capacité à travailler sous excel pour calculer des indicateurs et construire des graphiques associés.

Références bibliographiques

- Henri Leridon, les théories de la fécondité, Ined éditions, 2014
- Organisation internationale pour les migrations, Etat de la migration dans le monde, 2013 et s
- Benoît Ferry, l'Afrique face à ses défis démographiques, Karthala, 2008
- Catherine Rollet, Introduction à la démographie, 3e édition, Armand Colin, 2011
- Gabriel Poulalion et Georges Pupion, Analyse démographique, L'Harmattan, 2009

Langue d'enseignement

Français

Analyse textuelle

TD/TP : 12h

Enseignants : Elodie Baril(Ined)

Correspondants : Brigitte Gelein

Objectif de la matière

Mettre en œuvre les méthodes de la statistique textuelle sur des corpus de nature différente (questions ouvertes, mots associés, articles de presse, pages Web, entretiens, etc.) à l'aide de logiciels spécifiques (IraMuteQ, Spad, package tm, plugin de RCommander (R.TeMiS))
Interpréter et présenter les résultats, et adapter leur restitution à la demande.

Contenu de la matière

Retour sur l'origine et le développement des méthodes de la statistique textuelle.
Place de la statistique textuelle dans la mise en œuvre du Text Mining.
Présentation de différents types de corpus de textes, leur collecte et leur mise en forme.
Différentes étapes de traitement d'un corpus de texte : réduction du vocabulaire, **construction du lexique associé** (lemmatisation), les différents **tableaux lexicaux** et leurs traitements statistiques.
Résultats, **aides à l'interprétation** : vocabulaire spécifique, contexte d'utilisation des mots, les sorties des **analyses multivariées ou classifications et aide à la post-codification** d'une question ouverte.
Mise en œuvre d'une analyse au moyen de plusieurs logiciels : IraMuTeQ (méthode Alceste), Packages spécifiques de R comme R.TeMiS), Spad.
Restitution des résultats.

Pré-requis

Connaissance de base en statistique descriptive et en analyse des données (Analyse factorielle et classification).

Références bibliographiques

Lebart, L. et Salem, A. (1994), *Statistique textuelle*. Paris, Dunod, 342 p.
Garnier B., Guérin-Pace F. 2010. *Appliquer les méthodes de la statistique textuelle*. Paris, CEPED, 86 p. (*Les Clefs pour*) (<http://www.ceppe.org/?Appliquer-les-methodes-de-la>)
Guérin-Pace F., Saint-Julien T., 2012 : « *Les mots de L'Espace Géographique. Une analyse lexicale des titres et mots-clés de 1972 à 2010* », L'Espace Géographique, Tome 41, n°1, pp.4-30.
Brennetot A., Emsellem K., Guérin-Pace F., Garnier B. 2013. *Dire l'Europe à travers le monde. Les mots des étudiants dans l'enquête EuroBroadMap*, Cybergéo <http://cybergegeo.revues.org/25684>
Baril É., Guérin-Pace F. 2016. Compétences à l'écrit des adultes et événements marquants de l'enfance : le traitement de l'enquête Information et vie quotidienne à l'aide des méthodes de la statistique textuelle. In: Economie et statistique, n°490, Compétences des adultes à l'écrit et en calcul. pp. 17-36. www.persee.fr/doc/estat_0336-1454_2016_num_490_1_10719

Langue d'enseignement

Français

Entreprise

TD/TP : 12h

Enseignants : Florence Busson-Villa (Université de Rennes 2)

Correspondants : Brigitte Gelein

Objectif de la matière

Mettre en évidence et analyser les modes de fonctionnement des entreprises

Contenu de la matière

Introduction – Les rôles et classifications des organisations

La formulation de la stratégie

Le management des organisations

Les décisions financières

Conclusion – La performance et le climat social

Pré-requis

Aucun

Références bibliographiques

Autissier D., Bensebaa F., Boudier F. (2012), *L'atlas du management, Les meilleures pratiques et tendances pour actualiser vos compétences - 100 dossiers clés*, Eyrolles.

Barabel M., Meier O. (2015), *Manageor*, Dunod.

Brilman J. (2011), *Les meilleures pratiques de management*, Editions d'Organisation.

Denis J.P. et al (2016), *Lexique de gestion et de management*, Dunod.

Robbins S., DeCenzo D., Coulter M., Rüling C. (2017), *Management - L'essentiel des concepts et des pratiques*, Pearson Education.

Saussois J.M. (1998), *Comprendre les organisations, Sciences Humaines*, Hors Série n°20.

Théorie des indices de prix

Cours : 12h

Enseignant : Olga Vasyechko

Correspondant : Didier Le Bozec

Objectif de la matière

L'objectif de ce cours est la présentation de la théorie des indices de prix et son calcul dans la pratique d'indices officiels de la statistique publique.

Contenu de la matière

1. Généralités

- 1.1 Les objectifs de l'indice de prix
- 1.2 La famille des indices de prix
- 1.3 L'utilisation des indices de prix
- 1.4 Les règlements Européens

2. Les approches théoriques et les propriétés des indices

- 2.1 Les indices classiques
- 2.2 L'approche comptable
- 2.3 L'approche axiomatique
- 2.4 L'approche stochastique
- 2.5 L'approche économique

3. Les approches pratiques

- 3.1 Le cadre statistique des indices de prix
- 3.2. La construction des indices
- 3.3. Les séries temporelles et le chaînage

Prérequis

Une bonne maîtrise des concepts économiques et des outils statistiques.

Références bibliographiques

1. Compendium of HICP reference documents. European Commission. Luxembourg: Publications Office of the European Union, 2013.

2. Handbook on industrial producer price indices (PPI). Eurostat Methodologies and Working Paper, Luxembourg: Publications Office of the European Union, 2012.
3. Berthier Jean-Pierre, Introduction à la pratique des indices statistiques, INSEE, N°M0503
4. Chauvet-Peyrard Axelle, Les indices de prix. De la théorie à la pratique, 2014
5. DIEWERT Erwin, « On the stochastic approach to index numbers », groupe d'Ottawa, 1995
6. DIEWERT Erwin, HERAVI Saeed, SILVER Mick, « Hedonic Imputation versus Time Dummy Hedonic Indexes », groupe d'Ottawa, 2007
7. GUÉDÈS Dominique, « Fashion and consumer price index », groupe d'Ottawa, 2007
8. Reis Ricardo, A dynamic measure of inflation, Columbia University, 2009
9. Sillard Patrick, Wilner Lionel, Indices de prix à utilité constante et substitutions intermensuelles
10. SZULC Bohdan, « Linking price index numbers », paru dans Price Level Measurement: Proceedings of a conference sponsored by Statistics Canada, 1983
11. TRIPLETT Jack, « Should the cost-of-living index provide the conceptual framework for a consumer price index », groupe d'Ottawa, 1999
12. VIGLINO Lionel, « Le concept unificateur des indices de prix et proposition d'un nouvel indice », Journées de Méthodologie Statistiques, 2000
13. Indices de prix. Documentation méthodologique. INSEE

Langue d'enseignement

Français

UE 8 – Professionnalisation

Évaluation d'impact des Politiques Publiques : des exemples pratiques

Cours : 12h

Enseignants : Intervenants professionnels

Correspondants : Catherine Benjamin (Université de Rennes 1), Brigitte Gelein

Objectif de la matière

Les politiques publiques sont généralement mises en places pour améliorer des indicateurs de résultats, par exemple pour augmenter les revenus, faciliter l'apprentissage ou améliorer le niveau de santé. Savoir si les modifications espérées se sont effectivement produites est une question de politique publique fondamentale. Les évaluations d'impact visent ainsi principalement à déterminer si un programme a atteint ou non les résultats souhaités.

L'évaluation constitue aujourd'hui un enjeu majeur pour améliorer les politiques publiques et s'est ainsi imposée comme un domaine à part entière des sciences sociales. Réalisés principalement par des chercheurs, ces travaux reposent sur des méthodes statistiques capables de mesurer ce qui, dans les résultats observés, est réellement attribuable à la politique publique.

En France, de nombreux acteurs sont chargés d'une telle mission.

Cette UE propose de faire appel à des différents professionnels (France stratégie, Crest,) pour présenter des travaux réalisés sur différentes thématiques concrètes en matière d'évaluations.

Contenu de la matière

Interventions de professionnels sur différentes thématiques en matière d'évaluation

Pré-requis

UE Méthodes d'évaluation, UE Etude de cas

Langue d'enseignement

Français

Régression pénalisée

Cours : 12h

Enseignant : Denys Pommeret

Correspondant : Brigitte Gelein

Objectif de la matière

- Maîtriser les régressions Ridge, LASSO, Elastic Net et les variantes du LASSO.
- Savoir sélectionner les meilleures variables explicatives suivant différents critères.
- Maîtriser les package R associés.

Contenu de la matière

- 1/ Rappel sur les modèles linéaires généralisés.
- 2/ Extension à la régression pénalisée.
- 3/ Les méthodes Ridge, LASSO et ElasticNet.
- 4/ Les variantes du LASSO.
- 5/ Le cas des données longitudinales.
- 6/ Le cas des données manquantes.
- 7/ Régression par Machine Learning.

Pré-requis

Les participants devront avoir déjà étudié le modèle linéaire.

Références bibliographiques

Cf poly fourni.

Langue d'enseignement

Français

UE 9 – Stage

Stage

Correspondant : Brigitte Gelein

Pour les fonctionnaires en poste ou suivant le parcours STD (formation initiale ou continue), le stage est effectué dans le service d'affectation. Pour les élèves non fonctionnaires et fonctionnaires en formation initiale des parcours ES et MSP, le stage est effectué dans une entité dont l'activité entre dans le champ de compétences du master. Le sujet de stage sera soumis à l'approbation du responsable du master. Les étudiants devront rédiger également un rapport de stage qui fera l'objet d'une soutenance. Les modalités et les attendus de ce stage sont détaillés dans le guide des stages.

Majeure
**« Méthodologie de
la statistique publique »**

Séries temporelles avancées

Forecasting

Cours : 12h • TD/TP : 6h

Enseignant : Jean-Cyprien Héam (Drees)

Correspondant : Brigitte Gelein

Objectif de la matière

Maîtriser des techniques avancées de l'analyse des séries temporelles.

Contenu de la matière

Le cours est composé pour moitié de théorie et pour moitié de cas pratiques.

Le volet théorique couvre la modélisation à l'aide de variables exogènes (« étalonnage »), la modélisation par correction d'erreur, la notion de racine unitaire (variable « explosive ») et la notion de causalité.

Le volet pratique couvre l'estimation d'étalonnages, l'estimation de modèles à correction d'erreur, la détermination de la présence de racine unitaire, la détermination de la causalité d'une variable sur une autre, la construction de grilles délais et la problématique de l'estimation en temps réel. Audelà de la technique pure, une part importante de l'apprentissage est consacré à savoir prendre du recul sur les résultats obtenus et les méthodes employés. Les cas pratiques sont tous fondés sur des situations professionnelles authentiques (principalement de l'analyse conjoncturelle).

Les cas pratiques sont réalisés sous R.

Pré-requis

Connaissance de base sur les séries temporelles et les modèles univariés.
Bonne compréhension de la régression linéaire.

Contrôle des connaissances

Un examen écrit

Langue d'enseignement

Français

Modèles de prévision multivariés

Forecasting

Cours : 9h • TD/TP: 9h

Enseignant : Ludovic Gauvin

Correspondant : Brigitte Gelein

Objectif de la matière

L'objectif de cette UE est de présenter les techniques classiques de prévision des séries temporelles multivariées. Nous introduirons notamment les modèles autorégressifs vectoriels ainsi que les modèles à correction d'erreur et nous verrons comment ajuster et utiliser ces modèles en pratique.

Contenu de la matière

1. Processus VAR stationnaires

- 1.1 La notion de causalité
- 1.2 La modélisation VAR
- 1.3 Identification d'un modèle VAR et prévision
- 1.4 Causalité et variables macro-économiques

2. Cointégration

- 2.1 Notion de cointégration
- 2.2 Modèles VECM
- 2.3 Fonction de réponse impulsionnelle et décomposition de la variance

Pré-requis

Une bonne maîtrise des techniques générales de traitement des séries chronologiques

Contrôle des connaissances

Examen sur table (durée 1 h) 50%

Exercice pratique individuel de modélisation en salle informatique (durée 2h30) 50%

Références bibliographiques

Hamilton, "Time series Analysis", Princeton University Press
Gouriéroux, Monfort, "Séries temporelles et Modèles dynamiques", Economica
Lütkepohl, "New introduction to multiple time series analysis", Springer

Langue d'enseignement

Français

Enquêtes longitudinales

Cours : 12h • TD/TP : 6h

Enseignants : Guillaume Chauvet

Correspondant : Guillaume Chauvet

Objectif de la matière

Dans beaucoup d'enquêtes, l'objectif n'est pas seulement de produire des estimations ponctuelles dans le temps (ou transversales), mais également de produire des estimations en évolution (ou longitudinales) entre deux dates d'enquête. Dans le cas d'une enquête auprès des ménages, on va par exemple sélectionner un échantillon de ménages à la date t , et le suivre dans le temps en le complétant à la date $t+1$ avec un sous-échantillon additionnel de ménages, à la fois pour pouvoir représenter les unités nouvellement apparues et également pour compenser de l'attrition.

Cette méthodologie pose des problèmes d'estimation, notamment en transversal. D'une part, les ménages ne sont pas des unités stables dans le temps : un ménage peut se diviser, accueillir de nouveaux individus ou même disparaître. Pour produire une estimation à la date $t+1$, il faut pouvoir lier les ménages présents à ceux échantillonnés initialement. D'autre part, il faut incorporer dans l'estimation le sous-échantillon additionnel, qui permet en particulier de représenter les nouveaux ménages, sans faire de doubles comptes.

Ce cours propose une introduction à la méthode de partage des poids, qui constitue un outil simple et malléable pour permettre de produire des estimations transversales, notamment pour les enquêtes auprès des ménages.

Contenu de la matière

- 1) - Quelques rappels sur l'échantillonnage en population finie
- 2) - Le sondage indirect ou la méthode de partage des poids
- 3) - Enquêtes Longitudinales
 - a. Estimation longitudinale
 - b. Estimation transversale

Pré-requis

Cours de Théorie des Sondages

Contrôle des connaissances

Références bibliographiques

Ardilly, P. (2006). *Les techniques de sondage*. Editions Technip.

Lavallée, P. (2002). *Le Sondage Indirect: Ou la Méthode Généralisée Du Partage Des Poids*. Editions de l'Université de Bruxelles.

Langue d'enseignement

Français

Estimation sur petits domaines et compléments sur les redressements

Cours : 18h

Enseignant : Pascal Ardilly (Insee)

Correspondant : Brigitte Gelein

Objectif de la matière

Dans une première partie (2 jours), on s'intéresse aux problèmes d'estimation sur petits domaines. Quand les sous-populations d'intérêt sont de petite taille, la taille d'échantillon qui les recoupe est faible et conduit à des estimateurs très instables. Il est alors nécessaire de mobiliser toute l'information auxiliaire disponible, afin de proposer des méthodes d'estimation s'appuyant sur la théorie des sondages et sur une approche de modélisation.

Dans une seconde partie (1 jour), on aborde les techniques de redressement (dites aussi 'calage'). Un échantillon étant tiré, il s'agit de modifier à la marge le jeu des pondérations initiales afin de réduire significativement la variance d'échantillonnage, tout en conservant à très peu de chose près la propriété essentielle d'absence de biais. Cette méthodologie se décline dans différents contextes, en particulier il arrive que les conditions théoriques d'une mise en œuvre standard ne soient pas vérifiées.

Contenu de la matière

Partie 1 : Estimation sur petits domaines

- Estimation directe
- Estimation indirecte avec modélisation implicite
- Estimation indirecte avec modélisation explicite

Partie 2 : Redressements : rappels et compléments

- Rappels sur la théorie standard
- Application au cas des domaines
- Calage à plusieurs niveaux
- Contexte d'information auxiliaire obsolète
- Contexte d'information auxiliaire incertaine
- Contexte d'information auxiliaire peu ou non-explicative
- Calage et fonctions tronquées : choix des bornes

Pré-requis

Une bonne maîtrise des techniques de sondages telles qu'elles sont présentées dans le cours de Théorie des sondages, de Sondages Avancés et de Méthodologie d'Enquête.

Contrôle des connaissances

Examen ou Analyse de documents à rendre ultérieurement.

Références bibliographiques

- ARDILLY, P. (2006) , *Les Techniques de Sondage*, Paris, Technip.
- RAO, J.N.K., MOLINA I. (2015), *Small Area Estimation*, 2nd édition, New York, Wiley.
- SÄRNDAL, C-E., SWENSSON, B, WRETMAN, J. (1992), *Model Assisted Survey Sampling*, Springer-Verlag.

Langue d'enseignement

Français

Classification non supervisée et réduction de dimensions

Cours : 12h • TD/TP : 6h

Enseignant : Brigitte Gelein

Correspondant : Brigitte Gelein

Objectif de la matière

L'objectif de cette UE est de fournir aux étudiants des outils d'analyse factorielle et de classification non supervisée permettant de prendre en compte certains aspects de la structure des données. Les logiciels utilisés en TP seront les logiciels R et SAS.

L'Analyse Factorielle Multiple (AFM) est dédiée à l'étude des tableaux individus x variables dans lesquels les variables sont structurées en groupes (exemple : mêmes questions prises à différentes dates). Les variables actives peuvent être quantitatives et/ou qualitatives. A partir d'un tel tableau, l'AFM produit les résultats usuels de l'analyse factorielle et des résultats spécifiques tels que la représentation des individus du point de vue de chaque groupe, représentation des groupes etc. L'AFM Duale est, elle, dédiée aux tableaux individus x variables dans lesquels les individus sont structurés en groupes. On peut ainsi, par exemple, analyser l'évolution à différentes dates des corrélations entre variables. La complémentarité entre AFM et classification d'individus sera abordée avec la classification de trajectoires temporelles.

Par la création de groupes de variables corrélées, la classification de variables offre la possibilité de construire des représentants de classes (variables synthétiques) - ou encore de choisir parmi les variables initiales celles qui sont les plus représentatives de leur classe. Ce nouvel ensemble plus restreint de variables sera plus facilement gérable et interprétable.

Les Méthodes d'Appariement Optimal (Optimal Matching Analysis) permettent de construire une typologie de séquences, i.e. de suites d'éléments : étudier la succession des différents métiers qu'ont exercés des individus par exemple. La première étape de cette procédure consiste à calculer une distance entre les séquences. La seconde étape correspond à la classification proprement dite.

Contenu de la matière

1- Rappels succincts d'analyses factorielles simples

ACP (Analyse en composantes principales), ACM (Analyse des correspondances multiples).

2- Analyse conjointe de tableaux multiples

AFM (Analyse Factorielles Multiples) et AFM duale.

3-Classification de variables

Présentation de différents algorithmes selon la nature des variables à classer

4-Classification de trajectoires

Classification de courbes en lien avec l'AFM. Classification de séquences par Méthodes d'Appariement Optimal.

Pré-requis

Une bonne maîtrise des techniques générales d'analyses factorielles simples et de classification d'individus statistiques (classification ascendante hiérarchique, centres mobiles).

Contrôle des connaissances

Un quizz de 45 minutes + 1 compte-rendu de TP.

Références bibliographiques

Chavent, M, Kuentz, V. and Saracco J. (2009). A Partitioning Method for the clustering of Categorical variables. In *Classification as a Tool for Research*, Hermann Locarek-Junge, Claus Weihs (Eds), Springer, Proceedings of the IFCS'2009, Dresden

Chavent, M., Kuentz, V., Liquet B., Saracco, J., (2011). Classification de variables : le package ClustOfVar. *43 èmes Journées de Statistique (SFdS)*, Tunisie.

Collectif (1996), *L'analyse des données évolutives*, Technip.

Escofier B., Pagès J. (2008), *Analyses factorielles simples et multiples*, Dunod

Gabadinho, A., Ritschard, G., Müller, N.S. & Studer, M. (2011), Analyzing and visualizing state sequences in R with TraMineR, *Journal of Statistical Software*. Vol. 40(4), pp. 1-37.

Lesnard, L., De Saint Pol, T. (2006), Introduction aux méthodes d'appariement optimal (Optimal Matching Analysis), *Bulletin de méthodologie sociologique*, , Vol. 90, 5-25.

Pagès, J., *Analyse factorielle multiple avec R*, EDP sciences.

Studer, M., G. Ritschard, A. Gabadinho and N. S. Müller (2011). Discrepancy analysis of state sequences, *Sociological Methods and Research*, Vol. 40(3), 471-510.

Langue d'enseignement

Français

Multicolinéarité et régression PLS

Cours : 9h • TD/TP : 9h

Enseignant : Christian Derquenne (EDF)

Correspondant : Brigitte Gelein

Objectif de la matière

L'objectif de cette UE est de fournir aux étudiants des outils d'analyse des données et de régression qui permettent de mieux prendre en compte certains aspects de la structure des données.

Nous nous placerons plus précisément au sein de cette partie d'enseignement dans le cas où les variables peuvent être non seulement en nombre élevé, mais aussi être très liées de façon linéaire, par exemple. Ce qui peut poser un problème lorsque l'on désire modéliser un phénomène à l'aide de variables candidates à l'explication colinéaires. Les questions qui se posent de façon pratique sont alors les suivantes : Doit-on garder toutes les variables ? Que faire si le commanditaire de l'étude désire garder toutes les variables « explicatives » ? Dans ce cas, comment les pondérer, sans altérer l'interprétation métier du modèle et la signification statistique ? Par ailleurs, il peut arriver que le nombre de variables soit très élevé par rapport au nombre d'individus et que l'on désire garder, là encore, l'ensemble des variables. Comment construire un modèle qui réponde à ce souhait ? Le problème du grand nombre de variables peut également se poser pour le phénomène à expliquer. En effet, ce dernier peut très bien être mesuré sous différentes dimensions liées entre elles (plusieurs facettes). Dans ce cas, il y aura plusieurs variables réponses à modéliser de façon conjointe. Il ne serait alors pas adéquat d'ajuster celles-ci indépendamment. En d'autres termes, comment construire un modèle de régression multivarié pour les variables réponses ? Enfin, il arrive souvent que le tableau de données ne soit pas complètement renseigné, soit parce que certaines observations sont manquantes, soit parce que l'on a jugé que certaines de celles-ci sont atypiques (aberrantes) pour pouvoir de façon raisonnable les intégrer dans l'analyse statistique. Ce problème entraîne la question suivante : Comment élaborer un tel modèle, alors que la structure des données manquantes, même si celles-ci ne sont pas nombreuses pénalisent fortement l'estimation, en éliminant une très forte proportion des observations

Un certain nombre d'approches qui répondent à ces questions, permettent de contourner ces problèmes, mais de façon partielle. Les méthodes de sélection de variables explicatives permettent de garder les variables les plus influentes dans le modèle, mais pas toujours de façon efficace face à la multicolinéarité. Le modèle choisi peut alors être inintéressant sous l'aspect métier, voire incohérent. Par contre, la méthode de régression Ridge répond en partie à la question de prédicteurs liés entre eux au moyen d'une pondération de l'ensemble de ceux-ci, mais elle exige des paramètres non analytiques, donc difficiles à évaluer a priori. Le recours à la validation croisée est donc nécessaire. Par ailleurs, ces méthodes ne s'attaquent pas au problème de variables réponses multivariées, des données manquantes et d'un nombre important de variables explicatives face au nombre d'individus. La régression sur composantes principales, qui consiste à réaliser une ACP sur toutes les variables explicatives, permet de garder les composantes principales qui résument aux mieux les variables initiales. Enfin, la régression PLS (Partial Least Squares) répond à l'ensemble des questions posées précédemment : conservation de l'ensemble des variables explicatives, et corollairement prise en compte de la multicolinéarité, réponses multivariées, nombre élevé de variables face au nombre d'observations, données manquantes. En effet, cette

méthode repose sur le principe de l'algorithme NIPALS (Non linear Iterative Partial Least Squares) qui est très puissant en termes de robustesse. Cependant, même si cette méthode semble être satisfaisante, elle n'est pas parfaite, comme toutes les approches de modélisation.

L'objectif de ce cours est principalement de fournir des solutions aux différents problèmes soulevés précédemment, avec un point particulièrement développé sur la régression PLS et l'algorithme NIPALS. Ce cours est accompagné de différents cas réels permettant d'entrer progressivement de façon pédagogique dans l'étude de ces méthodes.

Contenu de la matière

1-Les différents problèmes abordés par les exemples

2-La multicollinéarité dans tous ses états

Décorticage de la multicollinéarité

Outils de détection de la multicollinéarité

Premières solutions pour pallier la multicollinéarité (régression sur composantes principales, régression sur typologie de variables)

3-L'algorithme NIPALS et les données manquantes

L'ACP sous l'angle de la déflation

L'algorithme NIPALS

4-La régression PLS

La régression PLS pour une réponse univariée (PLS1)

La régression PLS pour une réponse multivariée (PLS2)

Pré-requis

Une bonne maîtrise des techniques générales d'analyses factorielles simples et de classification de variables, ainsi que l'inférence dans le cadre du modèle linéaire gaussien.

Contrôle des connaissances

Compte-rendu de TP à rendre après le cours (un mois, environ)

Références bibliographiques

Tenenhaus (2007), *Statistique : méthode pour décrire, expliquer et prévoir*, Dunod

Tenenhaus (1998), *La régression PLS*, Technip

Apprentissage, agrégation de modèles et SVM

Cours : 12h • TD/TP : 6h

Enseignant : Brigitte Gelein

Correspondant : Brigitte Gelein

Objectif de la matière

Ce cours présente différentes méthodes d'apprentissage par agrégation de modèles. Les algorithmes décrits sont basés sur des stratégies adaptatives ou aléatoires. Ils permettent d'améliorer l'ajustement par une combinaison ou agrégation d'un grand nombre de modèles tout en évitant le sur-ajustement. Les Support Vector Machines (SVM) feront également l'objet d'une présentation.

La mise en application de ces méthodes est faite en atelier sous R principalement. L'objectif de cette UE est de fournir aux étudiants des outils de comparaison et sélection des modèles les plus performants dans la prévision d'une variable quantitative ou d'une variable qualitative.

Contenu de la matière

1. Bagging and random forest : agrégation par moyenne pondérée ou vote des estimations de différents modèles du même type, construits sur des sous-ensembles de la table d'apprentissage.
2. Boosting : agrégation par moyenne pondérée ou vote des estimations de différents modèles du même type construits sur toute la table d'apprentissage, chaque modèle est une version adaptée du précédent en donnant plus de poids, lors de l'estimation suivante, aux observations mal ajustées ou mal prédites lors de l'étape précédente.
3. Stacking. : agrégation de différents modèles de types différents par construction d'un modèle « superviseur » qui combine au mieux les prédictions des modèles primaires.
4. Support Vector Machines (Séparateurs à Vaste Marge ou SVM) : ils sont basés sur la recherche de l'hyperplan de marge optimale qui, lorsque c'est possible, classe ou sépare correctement les données tout en étant le plus éloigné possible de toutes les observations.

Pré-requis

Une bonne maîtrise des techniques générales d'analyses factorielles simples et de classification d'individus statistiques (classification ascendante hiérarchique, centres mobiles), ainsi que l'inférence dans le cadre du modèle linéaire gaussien.

Contrôle des connaissances

Un quizz de 45 minutes + un compte-rendu de TP.

Références bibliographiques

- P. BARTLETT and M. TRASKIN. AdaBoost is consistent, *JMLR*, 2007

- G. Biau, A. Fischer, B. Guedj et J. D. Malley, COBRA : *A Nonlinear Aggregation Strategy*, Journal of Multivariate Analysis 2013
- L. BREIMAN. Bagging predictors, Machine Learning, 24.2, 1996
- L. BREIMAN., *Stacked Regression*, Machine Learning, 24.1, 1996
- P. BUHLMANN, B. YU. *Boosting with the L2 loss: Regression and classification*, 98.462, 2003
- Y. FREUND and E. SCHAPIRE, *Experiments with a new boosting algorithm*, 1996
- R. GENUER. *Forêts aléatoires : aspects théoriques, sélection de variables et applications*, Thèse de l'Université Paris XI, 2010
- T. HASTIE, J. FRIEDMAN and R. TIBSHIRANI. *The elements of Statistical Learning: Data Mining, Inference and Prediction (2nd ed.)*, 2009

Langue d'enseignement

Français

Analyse spatiale

Cours : 9h • TP : 9h

Enseignant : Marie-Pierre de Bellefon (Insee)

Correspondant : Brigitte Gelein

Objectif de la matière

De plus en plus de données sont géolocalisées. L'ensemble des coordonnées géographiques des objets à traiter contient donc des informations potentiellement riches pour l'analyse. Que faire avec ces nouvelles sources de données ? Dans quels cas doit-on prendre en compte leur dimension spatiale ? Comment appliquer les méthodes de statistique et d'économétrie spatiale ?

Ce cours suit la démarche d'un analyste confronté à des données géolocalisées. A l'issue, les élèves sauront décrire et cartographier des données spatiales ; mesurer l'importance des effets spatiaux afin de juger de la nécessité de mettre en place une modélisation spécifique, et enfin choisir un modèle adapté aux données et aux objectifs de l'étude. Les zonages d'étude de l'Insee seront également présentés. L'ensemble des applications seront faites grâce au logiciel R.

Contenu de la matière

Décrire les données géolocalisées :

- Analyse spatiale descriptive : prise en main des données avec le logiciel R, réalisation de premières cartes, notions de sémiologie cartographique.
- Codification de la structure de voisinage : présentation de différentes notions de voisinage, fondées sur la contiguïté ou les distances entre observations. Enjeux liés à une bonne définition.
- Les zonages d'étude de l'Insee : aires urbaines, unités urbaines, grille de densité, zones d'emploi : méthode de constitution, interprétation, usages. La version présentée sera celle de 2010, avec des éléments sur la mise à jour prévue en janvier 2020.

Mesurer l'importance des effets spatiaux :

Les données spatiales peuvent être réparties en trois catégories : surfaciques, ponctuelles, continues. A chaque catégorie correspondent un ensemble de méthodes d'analyse.

- Indices d'autocorrélation spatiale : ces indices permettent de mesurer la force des interactions spatiales entre les observations. Les versions locales et globales des indices seront présentées.
- Configurations de points : cette partie donnera les méthodes et les outils permettant notamment de mettre en évidence les éventuelles attractions ou répulsions entre les différents types de points, et la façon dont on évalue la significativité des résultats obtenus.
- Géostatistique : des notions sur la façon dont on peut modéliser une variable en des points où elle n'a pas été échantillonnée seront abordées ici.

Prendre en compte les effets spatiaux

- Econométrie spatiale : ces modèles prennent en compte la dépendance spatiale, c'est à dire les situations où la valeur d'une observation est liée aux valeurs des observations voisines.

- Lissage spatial : le lissage permet de faire apparaître des phénomènes spatiaux qui resteraient inobservés si l'on cartographiait uniquement les objets bruts.
- Régression géographiquement pondérée : ces modèles répondent au constat qu'une régression estimée sur l'ensemble d'un territoire d'intérêt peut ne pas appréhender de façon adéquate les variations locales. On peut notamment repérer où les coefficients locaux s'écartent le plus des coefficients globaux, et construire les tests permettant d'apprécier si et comment un phénomène est non stationnaire.
- Économétrie spatiale sur données d'enquête : ce module présentera les écueils liés à l'estimation d'un modèle d'économétrie spatiale sur données échantillonnées, et évaluera les potentielles corrections proposées par la littérature empirique.

Pré-requis

Connaissance de base en économétrie et UE2 Économétrie avancé.

Contrôle des connaissances

Un examen écrit et un TP noté

Références bibliographiques

Anselin L. (2001) Spatial econometrics, dans : Baltagi B. (ed.), Companion to Econometrics, Basil Blackwell, Oxford.

Le Gallo J. (2004) Hétérogénéité spatiale, principes et méthodes, Economie et Prévision, vol. 162, pp. 151-172.

Langue d'enseignement

Français

Analyse de données d'enquêtes

Cours/TP : 18h

Enseignant : Brigitte Gelein (Ensaï), Ronan Le Saout (Ensaï)

Correspondant : Brigitte Gelein

Objectif de la matière

Mettre en œuvre les méthodes statistiques « classiques » sur des données issues d'enquêtes par sondage, en tenant compte des poids et du plan de sondage.

Contenu de la matière

Rappels en théorie des sondages

Estimation et précision de paramètres : totaux moyennes, proportions, ratios, quantiles

Tests d'adéquation à une loi discrète, tests d'indépendance

Modèle de régression linéaire

Modèle de régression logistique

Mise en œuvre informatique (SAS, Stata, R)

Pré-requis

Connaissances de base en théorie des sondages, en régression linéaire, en régression logistique.

Contrôle des connaissances

Un examen écrit.

Références bibliographiques

Chambers, R.L., & C.J. Skinner (ed) [2003], *Analysis of Survey Data*, Wiley

Davezies L. et X. D'Haultfoeuille, [2012], *Faut-il pondérer ? Ou l'éternelle question de l'économètre confronté à des données d'enquête*, Actes des Journées de méthodologie statistique, Insee

Le Guennec, J. [2005], *La régression sur échantillon avec SAS*, Actes des Journées de Méthodologie Statistique, Insee

Skinner C.J., D. Holt & T.M.F. Smith (ed.), [1989], *Analysis of complex surveys*, Wiley

Langue d'enseignement

Français

Majeure
**« Statistique et traitement
des données »**

Développement d'applications

Cours : 15h • TP : 33h

Enseignant : Mathieu Acher (Université Rennes 1), Karim Tachfint (Insee)

Correspondant : Roamric Gaudel

Objectif pédagogique

L'objectif de ce cours est d'introduire les moyens de concevoir des applications informatiques de qualité (répondant aux besoins, évolutives et faciles à maintenir). Il s'agit de présenter l'ingénierie dirigée par les modèles en positionnant la conception dans les cycles de développement, et en mettant l'accent sur les enjeux et les pièges à éviter.

Le cours présente une introduction aux modèles de conception classiques, base du génie logiciel autour des technologies objet, en proposant des applications pratiques au cours de travaux pratiques et en étudiant des patrons de conception développés en Java. L'objectif de cet enseignement est également d'apprendre à développer et déployer un site Web dynamique en Java. Ce cours permet de se familiariser avec les architectures n-tiers et les serveurs d'applications et de bien maîtriser les principaux outils et langages avancés de développement des applications Web/JavaEE.

Contenu de la matière

S.4.2- Génie logiciel

1. Le génie logiciel
 - Pourquoi ?
 - Comment ?
2. Introduction au génie logiciel et bonnes pratiques de conception.
3. Architecture logicielle et modèle en couche, Exemple sur GWT.
4. Principaux patrons de conception, principe et mise en œuvre en Java.
5. Le test logiciel et l'ingénierie des langages.
6. L'ingénierie dirigée par les modèles.

S.4.2- Développement d'application web

1. Les Technologies, L'architecture et Les applications JavaEE
2. Interfaces d'accès aux applications (API et IHM)
3. Structuration des composants applicatifs (contraintes et solutions)
4. La persistance des données avec JPA

Pré-requis

Notation UML, connaissance du langage JAVA.

Contrôle des connaissances

TP notés.

Références bibliographiques

- I. SOMMERVILLE, *Le Génie logiciel*, Addison Wesley-France, 1988
- B. BEIZER, *Software Testing Techniques*, Second Edition, Van Nostrand, 1990
- B.W. BOEHM, *Software Engineering Economics*, Prentice-Hall, 1981
- E. GAMMA, R. HELM, R. JOHNSON, J. VLISSIDES, *Design patterns, catalogue de modèles de conception réutilisables*, Vuibert, 2007

Langue d'enseignement
Français

Réseaux, systèmes d'exploitation, sécurité de l'information

Networks, systems and Information Security

Cours-TD-TP : 38h

Travail personnel : 15h

Enseignants : David Boucher (Orange Business Services)

Correspondant : Romaric Gaudel

Objectif de la matière

Cet enseignement vise à présenter les systèmes d'information sous deux angles :

- Un angle technique détaillant l'architecture intrinsèque des réseaux tcp/ip et l'architecture fondamentale des systèmes d'exploitation
- Un angle sécurité qui présente les mécanismes techniques et organisationnels qu'il convient de mettre à œuvre afin d'assurer la protection de l'information, en termes de disponibilité, confidentialité, intégrité, traçabilité.

Principaux acquis de la formation : à l'issue du cours, l'étudiant saura

- Comprendre les modèles architecturaux des systèmes d'informations, les erreurs à éviter et les bonnes pratiques lors de la conception d'une plateforme
- Identifier les problématiques de protection des données sensibles, les questions sécurité à adresser pendant tout le cycle de vie d'un projet sur un S.I.

Contenu de la matière

1. Architectures réseaux
 - 1.1. Introduction : classification, protocoles, modèles OSI et TCP/IP
 - 1.2. Couche Hôte-réseau : couche physique, liaison de données, ethernet, xDSL, 801.11
 - 1.3. Couche Internet : IP, adressage, routage, ARP
 - 1.4. Couche Transport : ports, TCP, UDP, ICMP
 - 1.5. Couche Application : client/serveur, DNS, SMTP, HTTP
 - 1.6. Protocoles connexes, NAT, VPN IPsec, VPN SSL, VOIP
 - 1.7. Revue d'architectures
2. Architecture des systèmes d'exploitation
 - 2.1. Les processus, la communication inter-processus et l'ordonnancement
 - 2.2. La gestion de la mémoire
 - 2.3. Les systèmes de fichiers
 - 2.4. La virtualisation
3. Cybersécurité
 - 3.1. Intro, notions de base
 - 3.1.1. Les enjeux de la sécurité des systèmes d'information
 - 3.1.2. Les besoins de sécurité
 - 3.1.3. Vulnérabilité, menace, attaque
 - 3.1.4. Panorama de quelques menaces
 - 3.1.5. Le droit des TIC
 - 3.2. Règles d'hygiène informatique. Les 42 mesures préconisées par l'ANSSI
 - 3.2.1. Sensibiliser et former
 - 3.2.2. Connaître le système d'information

- 3.2.3. Authentifier et contrôler les accès
- 3.2.4. Sécuriser les postes
- 3.2.5. Sécuriser le réseau
- 3.2.6. Sécuriser l'administration
- 3.2.7. Gérer le nomadisme
- 3.2.8. Maintenir le système d'information à jour
- 3.2.9. Superviser, auditer, réagir

Des Travaux Pratiques complètent l'enseignement, comprenant notamment :

- La mise en œuvre d'un réseau local et la configuration des terminaux
- Des captures de trafic réseau (HTTP, SMTP) à des fins d'analyse et de récupération d'informations
- L'étude de processus sous Linux et Windows
- L'analyse de média USB
- L'étude des composants majeurs d'une carte-mère
- Exercices de mise en situation avec analyses de scénarios de protection d'un système d'information

Contrôle des connaissances

Un examen écrit.

Références bibliographiques

- Réseaux. Andrew Tanenbaum & David Wetherall. Pearson Education. 5^e édition. ISBN-10: 2-7440-7521-3
- Modern Operating Systems. Andrew Tanenbaum. Pearson Education. 3^e édition. ISBN-10: 1292038195.
- Publications de l'ANSSI. <https://www.ssi.gouv.fr/guide/guide-dhygiene-informatique/>

Langue d'enseignement

Français

Echanges sur le Web

Cours : 10h • TD : 10h • TP : 28h

Enseignant : Olivier Chantrel (Orange), Karim Tachfint (Insee)

Correspondant : Romaric Gaudel

Objectif du module « Les éléments de base d'une application web »

Ce cours vise à donner aux étudiants une vision détaillée du web d'aujourd'hui en présentant les technologies historiques du web (html, xml) et les technologies plus récentes du web 2.0 (css, javascript, HTML5, Ajax, Php, MySql). Ce cours introduit également les concepts de mise en œuvre des systèmes d'information d'aujourd'hui, basé sur les technologies web. Nous verrons notamment comment architecturer, concevoir et mettre en œuvre des applications, les différents moyens d'intégration permettant de faire communiquer ces applications entre elles.

Les différents concepts seront mis en pratique sur des technologies java et Spring.

Principaux acquis de la formation : à l'issue du cours, l'étudiant saura

- Avoir un regard critique sur les différents types d'architectures web
- Concevoir des applications basées sur les technologies web
- Définir et mettre en pratique des techniques et API d'intégration d'applications

Contenu de la matière

S.6.1- Les éléments de base d'une application web

1. L'historique du web / XML et ses applications (html, xml, dtd, web services etc.)
2. Le web actuel (css, javascript, HTML5, Ajax, MySql, Php)

S.6.2- Communication entre application web

- 2.1. Architecture et principes d'intégration
- 2.2. Web services
- 2.3. Web sémantique

Pré-requis

Connaissance des concepts Objets et du langage Java ainsi que du langage HTML.

Contrôle des connaissances

Un examen écrit

Références bibliographiques

- E. R. HAROLD, *The XML Bible*, IDG Books
- J.. BERNADAC, F. KNAB, *XML et Java : Avec trois études de cas*
- A. MICHARD, *XML, Langage et applications*, Eyrolles
- M. KEY, *XSLT Programmer's Reference*, 2nd Edition, Wrox Press

Masses de données distribuées

Datawarehouse and databases

Cours : 10h • TP : 38h

Enseignant : Shadi Ibrahim (Inria-Rennes), David Gross-Amblard (Irisa)

Correspondant : Romaric Gaudel

Objectif de la matière

Cet enseignement se décompose en deux parties spécifiques du domaine Big Data. La première concerne les architectures et la seconde les bases de données.

La première partie de cet enseignement vise à donner aux étudiants les connaissances de base sur les architectures distribuées spécialisées dans le traitement du Big Data. Les hypothèses, concepts et algorithmes seront présentés pour chacune d'entre elles. L'objectif est d'avoir une connaissance des systèmes sur Cloud disponibles actuellement et de pointer les directions futures de ces architectures. L'objectif final est la mise en œuvre avec Hadoop et une base de données Big Data spécifique.

La seconde partie a pour objectif de présenter les différents systèmes de gestion de bases de données qui sont utilisés aujourd'hui dans le domaine du Big Data. Nous présentons ici les quatre principaux types de bases appelées communément bases NoSQL : « clé-valeur », « colonne », « document » et « graphe ». Un exercice pratique complet sera fait sur avec la base de données de type documents, MongoDB.

Contenu de la matière

- Architectures Big Data
 - Introduction aux infrastructures distribuées : clusters, supercalculateurs, grilles, clouds
 - Introduction au cloud computing
 - Big Data: introduction, défis, enjeux
 - Explicit Data management
 - Transparent Data Management: NFS, Gfarm, Google File System:
 - Introduction à MapReduce et Hadoop:
 - Atelier pratique sur Hadoop
 - Avenir de MapReduce: défis
 - Approches post-MapReduce: Shark/Spark (Berkeley)
 - Approches post-MapReduce: Stratosphere
- Bases de données NoSQL
 - Les limites du modèle relationnel
 - Les apports des modèles NoSQL
 - Les différents types de bases
 - Type « clé-valeur »
 - Type « colonne »
 - Type « document »
 - Type « graphe »
 - Mise en oeuvre avec MongoDB et Cassandra

Pré-requis

Système réparti, interrogation (SQL) de bases de données relationnelles.

Contrôle des connaissances

Deux TP, un pour chacune des parties du cours.

Références bibliographiques

- G. PLOUIN, Cloud computing et SaaS, Editions Dunod
- Le livre blanc du Cloud, du SaaS et des Managed Services pour les partenaires IT et télécoms. Edition 2013
- R. HENNION, H. TOURNIER, E. BOURGEOIS, Cloud computing : Décider - Concevoir - Piloter - Améliorer, Editions Eyrolles, 2012

Langue d'enseignement

Français

Parcours
« Etudes statistiques »

Evaluation des politiques publiques

Cours : 40 h

Enseignant : Catherine Benjamin, Laurent Denant-Boemont, David Masclat (Université Rennes 1)

Correspondant : Catherine Benjamin

Objectif de la matière

Présentation des principales approches d'évaluation de politiques publiques : analyse coût-bénéfice et les outils d'évaluation d'impacts (à partir de données non expérimentales ou de données issues d'expériences randomisées, ou d'expériences en laboratoire, etc.).

Contenu de la matière

Le cours s'attachera à exposer les problèmes posés par l'évaluation des politiques publiques et les différentes approches qui ont été développés. Les principaux concepts seront étudiés. Les méthodes d'évaluation quantitative se différencient par la nature des données utilisées. La première famille de méthodes est celle des évaluations qui utilisent des données d'expériences contrôlées qu'il s'agisse d'expériences de terrain (les évaluations aléatoires) ou d'expériences contrôlées en laboratoire. La deuxième famille correspond aux évaluations qui mobilisent des données non expérimentales mais qui tentent de se rapprocher des conditions des données expérimentales. Ce cours propose une introduction à ces outils, en insistant sur les détails de leur implémentation pratique sous R. Une partie spécifique portera sur l'analyse coût-bénéfice, aussi appelée coût-avantage, qui correspond à une méthode d'aide à la décision qui permet d'évaluer les avantages et bénéfices d'un projet.

Pré-requis

Concepts fondamentaux de l'analyse microéconomique (surplus), Méthodes d'estimation d'un modèle de régression linéaire multiple

Contrôle des connaissances

Examen final écrit

Références bibliographiques

- Bruno Crépon et Nicolas Jacquemet, *Économétrie : méthodes et applications*, De Boeck, 2010
- Pauline Givord, 2010, *Méthodes économétriques pour l'évaluation de politiques publiques*, document de travail du D3e, Insee, G2010/08 2010 (http://www.insee.fr/fr/publications-et-services/docs_doc_travail/G2010-08.pdf)
- Marie Claire Villeval « L'Économie comportementale du marché du travail » (Les presses de Sciences Po)
- Nicolas Eber en Marc Willinger *L'économie expérimentale* » collection repère la découverte

Langue d'enseignement

Français

Etudes de cas de politique publique

Cours : 40h

Enseignant : Marie-Hélène Hubert, David Masplet, Laurent Denant-Boemont (Université Rennes 1)

Correspondant : Marie-Hélène Hubert

Objectif de la matière

Présentation de thèmes liés à l'évaluation de politiques publiques. Le cours prévu est divisé en 3 modules thématiques : mesures des consentements à payer, puis des études de cas appliquées au marché du travail, à l'économie de l'environnement.

Contenu de la matière

Mécanismes expérimentaux d'élicitation des dispositions à payer

Depuis la catastrophe de l'Exxon Valdez en 1989 aux USA, la question de la mesure économique des dommages environnementaux est devenue un enjeu fondamental des politiques publiques et des arbitrages juridiques. La méthode la plus souvent mise en avant pour mesurer la valeur subjective des dommages pour les individus est l'évaluation contingente. Cette méthode fait toutefois l'objet de nombreuses critiques. L'objectif de ce module est d'initier les étudiants à la mesure économique des dispositions à payer ou des dispositions à accepter afin d'identifier les principaux biais des méthodes possibles. Un focus particulier est fait sur les méthodes d'expérimentation en laboratoire qui permettent de contrôler certains biais possibles dans les mesures. Spécialement, les mécanismes d'enchères seront mis en avant parmi d'autres mécanismes de révélation possibles. Le module se base sur une série d'applications expérimentales ou sur données de terrain.

Marché du travail

Après avoir présenté le fonctionnement du marché du travail (demande et offre de travail), nous étudierons les politiques de l'emploi visant à réduire le chômage. En particulier, nous présenterons les principaux résultats des évaluations randomisées relatives à l'appariement des demandeurs d'emploi et des emplois vacants ainsi que les politiques visant à améliorer les qualifications des demandeurs d'emploi.

Economie de l'environnement :

Deux grands thèmes d'actualité en économie de l'environnement : le premier thème s'intéresse à la fiscalité écologique et au marché des droits à polluer. Quant au deuxième thème, il aborde la question de la croissance durable. Pour chaque thème, les outils de la théorie économique sont exposés. Puis, des applications économétriques et/ou empiriques sont étudiées.

Pré-requis

- Principes fondamentaux en microéconomie et de macroéconomie - Outils d'évaluation d'impacts (UE 5)

Contrôle des connaissances

Projet et Evaluation écrite

Références bibliographiques

- SCHUBERT, K., Paul Z.. *L'environnement: une nouvelle dimension de l'analyse économique*. Vuibert, 1998.
- BORJAS (2012) *Labor Economics*" McGraw Hill Edition
- CAHUC P. et ZYLBERBERG A. (2004), *Labor Economics*, MIT Press, Cambridge (2001, *L'économie du travail*, chez De Boeck)
- LESUEUR ET SABATIER (2008) *Microéconomie de l'emploi :Théories et applications*" (De Boeck)
- Lusk, J. and Shogren J.F. (2007). *Experimental auctions*, Cambridge University Press.
- POLITIQUES DE L'EMPLOI APPRENDRE DE L'EXPERIENCE, Bruno Crépon et Gerard Van den Berg (Science Po Presses)
- Schubert, Katheline, and Paul Zagamé. *L'environnement: une nouvelle dimension de l'analyse économique*. Vuibert, 1998.
- Weil, David. *Economic Growth*. Pearson, Third Edition, 2013.
- Cherry T., Kroll S. and Shogren J.H. (2008). *Environmental economics, experimental methods*.
- Weil, David. *Economic Growth*. Pearson, Third Edition, 2013.

Langue d'enseignement

Français

E3 – Majeure ES

Analyse de la conjoncture

Cours : 40h

Enseignant : Catherine Benjamin, Nathalie Payelle (Université Rennes 1), Janick Bérel (Conseil régional de Bretagne)

Correspondant : Catherine Benjamin (Université Rennes 1)

Objectif de la matière

- Analyser les facteurs explicatifs de la conjoncture macroéconomique et sectorielle
- Comprendre la formation des prix des matières premières et leur impact sur la conjoncture
- Savoir rechercher les informations pour établir des notes de conjoncture
- Analyser l'impact de l'environnement international

Contenu de la matière

Le cours est divisé entre 4 modules thématiques :

- **Analyse de la conjoncture régionale :**

Ce module a pour objectif

- de présenter les différents types d'indicateurs de la conjoncture économique régionale et les méthodologies employées,
- de déterminer les sources d'information disponibles,
- de cerner les attentes d'une note de conjoncture régionale,
- de présenter l'usage de ces notes en termes de politique économique.

- **Fonctionnement des marchés de matières premières :**

Ce module porte sur les enjeux économiques et environnementaux des marchés des matières premières. Ces produits étant des intrants indispensables de l'économie réelle, leur évolution de prix a des conséquences directes et immédiates sur les investissements, la production ou encore l'alimentation. Il s'agira d'étudier la formation des prix de ces produits issus de ressources naturelles et l'impact du caractère épuisable de la ressource, les facteurs explicatifs de la volatilité des prix et le rôle de la financiarisation des marchés. L'analyse économique du marché du pétrole fera l'objet d'un cours spécifique.

- **Analyse de notes de conjoncture macroéconomique :**

- Comprendre comment sont structurées les notes de conjonctures nationales et internationales
- Analyse des différents moteurs de la croissance de court terme
- Etude des scénarios de prévision à court terme

- **Environnement international et conjoncture.**

Il s'agira de d'étudier quels sont les possibles impacts de l'environnement international sur la conjoncture et notamment les plusieurs sources d'incertitude. On s'intéressera notamment à l'impact de la guerre commerciale à laquelle se livrent les deux premières

puissances mondiales, les Etats-Unis et la Chine, aux incertitudes grandissantes autour du Brexit et à la mesure des impacts des accords commerciaux.

Pré-requis

Connaissance de base en économétrie, en modélisation en équilibre partiel, et des grands principes macroéconomiques

Contrôle des connaissances

Deux projets sur 2 thématiques, sujets distribués la première semaine de cours

Références bibliographiques

- Notes de conjoncture des différents organismes nationaux et internationaux
- GIRAUD P-A (2006) L'économie des matières premières
- KRUGMANK., OBSTFELD M., et al. (2018) Economie internationale, Paris, Pearson Education

Langue d'enseignement

Français