

# Fouille de données

NGUYỄN Thị Minh Huyền

Email: [huyenntm@vnu.edu.vn](mailto:huyenntm@vnu.edu.vn)

VNU University of Science, Hanoi

# Plan

1. Introduction

2. Plan du cours

# Plan

1. Introduction

2. Plan du cours

# Plan

1. Introduction

2. Plan du cours

# Responsable du cours

- NGUYEN Thi Minh Huyen
- Département de Mathématiques, de Mécanique et d'Informatique  
Université des Sciences (Université nationale de Hanoi)
- Domaine de recherche : Ressources pour le traitement automatique des langues

# Introduction du cours

- Fouille de données : techniques et méthodes pour l'extraction automatique d'informations ou de connaissances nouvelles ou cachées dans des collections de données (qui peuvent être de très grande taille et rassembler différents types de média).
- Fouille de textes : Recherche et extraction d'information
- Matériels et références du cours sont fournis via la classe Google : <https://classroom.google.com/c/MzU2NzE4NjY2MTU4?cjc=7tae714>

# Plan

1. Introduction

2. Plan du cours

# Plan du cours

## Cours

- Introduction à la fouille de données (*Data Mining*) et à l'analyse de données textuelles (*Text Mining*)
- Rappel de probabilités
- Exploration des données : cas d'une et de deux dimensions
- Exploration des données multidimensionnelles - Apprentissage non supervisé
- Exploration des données multidimensionnelles - Apprentissage supervisé
- Introduction à l'analyse de données textuelles



# Plan du cours

## TPs

- Exercices avec logiciel à faire en binôme
- Exercices individuels
- Projet final en binôme
- Données (cf. Références)

Note finale : Exercices (40%) + Projet final (60%)