



**POLYTECHNIQUE
MONTRÉAL**

**LE GÉNIE
EN PREMIÈRE CLASSE**

David TREMBLAY 1748125 GROUPE 3

Billy BOUCHARD 1850477

Francisco CANELA GUTIERREZ 1878175

Travail de session 1/2

Présenté à Maha Gmira

MTH2302D- Probabilités et statistiques



Groupe 01



Département de mathématiques et génie industriel

Le 3 mars 2017

École polytechnique de Montréal

1 Contexte général des données

Le football américain est l'un des sports de contact les plus compétitifs des États-Unis et est actuellement pratiqué dans plus de 60 pays. Une équipe de football américain est composée d'environ 45 joueurs. Il existe deux premières équipes, l'équipe offensive et l'équipe défensive, composées par 11 joueurs chacune. En parlant de l'équipe offensive, parmi les 11 joueurs, on trouve les receveurs de passe. Fondamentalement, ils sont responsables de recevoir des passes du quart-arrière et normalement ils sont situés dans les côtés du terrain. Leur travail dans l'équipe est considéré très important pour obtenir des touchés (qui est en fait le principal objectif du jeu) et donc leur rendement sur le jeu est constamment observé et mesuré par la même organisation qui régleme ce jeu (NFL). La qualité de leur travail n'est pas seulement évaluée par le nombre de réceptions accumulées, mais aussi par les verges parcourues par réception. En général, les données qui seront présentées dans ce rapport décrivent le rendement des cent meilleurs joueurs de la saison passée (2016). Ce rendement est composé de plusieurs variables comme le nombre de réceptions, nombre des verges, verges/partie, nombre de touchdowns, pourcentage de réception, parmi des autres.



2 Provenance des données

Les données qui seront présentées ont été prises de la section des statistiques du site web de la NFL (National Football League) et sont disponible sur le site <http://www.nfl.com/stats/categorystats?tabSeq=0&statisticCategory=RECEIVING&season=2016&seasonType=REG>. L'expérience varie dans chaque variable qu'on trouve dans les données. Par exemple, en parlant de la variable aléatoire « nombre des réceptions », l'expérience est « le joueur a attrapé le ballon ». En parlant de la variable aléatoire « nombre de touchdowns », l'expérience est « le joueur a fait un touchdown ». Toutes les variables pour chaque joueur qui ont été mises sur le site web ont été obtenues par simple observation de chaque match. Il s'agit d'un processus très rigoureux duquel toutes les statistiques de la NFL sont produites.



3 Description de la forme des données

Cent joueurs de la NFL sont répertoriés dans notre grille *Excel*. Mis à part leur nom et leur équipe, dix autres variables représentant diverses statistiques y sont présentées pour chacun des joueurs. Deux de ces variables sont continues et les huit autres sont discrètes. Les noms des variables rendent la description des unités de mesure inutiles puisqu'elles sont évidentes à déduire. Par souci de clarté, mentionnons toutefois que la variable "plus longue réception" est en verges comme toutes les variables qui ont rapport à des distances puisque c'est l'unité utilisée au football américain.



Équipe: C'est une variable discrète qui fournit l'abréviation de l'équipe dans lequel le joueur progresse (IND pour Indianapolis par exemple).

Réception: C'est une variable discrète allant de 0 à l'infini qui représente le nombre de réceptions effectué par le joueur dans la saison.



Nombre de Verges: C'est une variable continue infinie positive qui représente le total de verge réussi par le joueur. (Les valeurs sont arrondies à l'entier.)



Verge/partie: C'est une variable continue et infinie positive représentant le la moyenne du nombre de verges parcourues par partie.



Plus longue réception: C'est une variable infinie positive dénombrable qui représente le nombre de verges pour la plus longue passe reçu par le joueur.



Nombre de touché: C'est une variable infinie positive dénombrable qui représente le nombre de touché effectué par le joueur. Un touché est quand le ballon est transporté jusque derrière la ligne de but de l'équipe adverse.



Réceptions de plus de 20 verges: C'est une variable infinie positive dénombrable qui représente le nombre de réceptions de passes qui étaient de 20 verges ou plus de distances.



Réceptions de plus de 40 verges: C'est une variable discrète **infinie** positive **dénombrable** qui représente le nombre de réceptions de passes qui étaient de 40 verges ou plus de distances.



Réception pour un 1er essai: c'est une variables discrète **infinie** positive qui représente le nombre de fois que le joueur à donner un premier essai a son équipe (franchi la ligne de 10 verges)



Pourcentage de réception menant à un premier essai: il s'agit d'une variable continue comprise entre 0 et 1 qui dit **combien** des réceptions ont permis l'obtention d'un premier essai.





Ballon échappé: C'est une variable **infinie** positive **dénombrable** qui représente le nombre de fois que le joueur a échappé le ballon (Il a capté la passe mais l'a échappé ensuite).



3.1 Questions ouvertes

Cette section regroupe quelques questions qui nous viennent à l'esprit lorsqu'on regarde notre fichier de données. Ces questions ne seront toutefois pas répondues dans ce rapport, mais plutôt dans le deuxième travail de session.

- Existe-t-il une relation entre le nombre de verges par partie et le nombre de touché effectué? 
- Quel est la variable qui influence positivement le plus les autres variables ? 
- Le nombre total de verges effectuées lors de la saison a-t-il un lien avec l'équipe du joueur?