Using data science & innovation to build a sustainable tertiary education model

C. Gans Combe – Senior researcher INSEEC
**FROM DEFORM**

Deform intended to provide potential symptoms of potential lack of integrity as well as a value-based model. In the end, how to disseminate those adequately?

- A professional, lifelong learning certificate
- Not targeted at academia
- An open source tool (freely available using EPALE)
- Not an “insider” job but trying to go virale

**TO PRORES**

The question of the lack of trust in: science? scientists? academia?

- Trust in the ones providing the knowledge...
- Not a question of lack of trust in science (70/30 rate for 50 years)
- A question of answering the questions asked...
- Not a question of topic, A Question of making matches: challenges are the same, whatever the field?
- More than delivering knowledge...
1. High level of enrollment – low level of achievement
   - A question of misguided orientation?
   - A question of unsuitability for employment?
   - A question of Trust? confidence in the institution to lead everyone towards the goal they have set for themselves: employment? Values?

2. High costs for the Community & for which impact on growth... A10 Indicator
   - A dichotomy of expectations, a misunderstanding of results: appetite/efficiency trade-off.
   - The wrong choice of regressors? => biased questioning
   - Alarming Facts: Spendings multiplied by 7 – for a limited impact on growth (spendings 2 time higher than the impact)

3. Why such a result => the cost of dropouts? Can education efficiency be enhanced?

Sustainability => rational use of resources including in education
TESTABLE PREDICTIONS
Building educational processes to match the F1 & F2 Factors
=> reduce dropouts

DATA TO TEST PREDICTIONS
OECD & EUROSTAT Samples
Expressions of the social personae

HYPOTHESES
=> We have the data to fit the persona
=> We have the data to built the theoretical persona targeted by specific curriculum

THEORY
By taking into consideration the weight of the F1 & F2 factors when building curriculum, we can increase students' adherence to these & their success rate, => direct impact on growth

QUESTIONS
=> Can this guidance be built?
=> Will it impact the dropout rate?

OBSERVATIONS
=> first choice curriculum: socially based (biased?)
=> Unfit social persona (ecosystem persona)

=> Guidance to discover social persona and the training that fill adequately fit this persona

HYPOTHESIS
=> Tertiary education is a growth factor, the less the tertiary education a population has, the less its growth is efficient. Direct cause/effect growth & educated workforce availability (OECD) => A10 indicator

=> If tertiary education is a growth factor, then dropout rates can be correlated to "losses" in the growth potential, as such the cost is not only what is spend, but also what is lost in growth potential

=> Why do students drop-out?
=> An indicator that can be cross educational sector?

PROBLEM
High enrollment => high dropout => high costs => high impact
What are the needed regressors?

STATEMENT
=> Dropout is correlated to two factors each including correlated subfactors:
=> F1 lack of trust in the academic sector to fulfill its task (pathway to personal equilibrium (social persona) - pleasure - free time - work - personal appreciation - financial comfort (Fulfilling function)
=> F2 lack of adequation of the educational orientation with one's expectations (difficulties to correct this when one pathway is taken => problematic of academic fields & silos/specialties (we want to minimize the
Testing our hypothesis required three phases: (1) collecting and defining the regressors - (2) sampling the (right) population - (3) Testing per -se

\[
F2 = P \left( \sum_{n} P_s - \sum_{n} P_c \right) = 0
\]

F1 will measure the proximity of the academic methodology and the social persona => trust equation (fuzzy logic - > link natural language expressions with a meaningful numerical analysis)
**Mismatch factors**

**What are** the mismatch factors (Patterns)?

**Question**
When does the mismatch occur?

**Achieving population**
- Match between the social persona & curricular persona

**Non-achieving population**
- Mismatch!

**Match level in non-achieving population (at a given time)**
- 14.2%
- Low level of match between the curriculum and social persona

**Match level in achieving population (at a given time)**
- 91.3%
- The mismatch factors (Patterns)
MISMATCH ANALYSIS

15% SELDOM CORRECTED
No change of Pattern => start to fail/inevitably fail

70% POTENTIALLY PRE-CURRICULAR
Compare the mismatch in time (appear very early in the time pattern)

31% ITERATIVE PATTERNS
A failure does not always imply a stop in the training process; but more a pause
CONCLUSION

Can we build an algorithm to reduce the mismatch level? What will be its impact?

In other words, typical inference question => what will an x level of dropout reduction have on:
=> Education efficiency
=> For what it takes, adhesion to educational programs

- We have the patterns
- We have the regressors
- Further data to answer the inference question & measure the statistical adequacy of this pathway...