

Visualisation de données: populations microbiennes issues des toilettes publiques



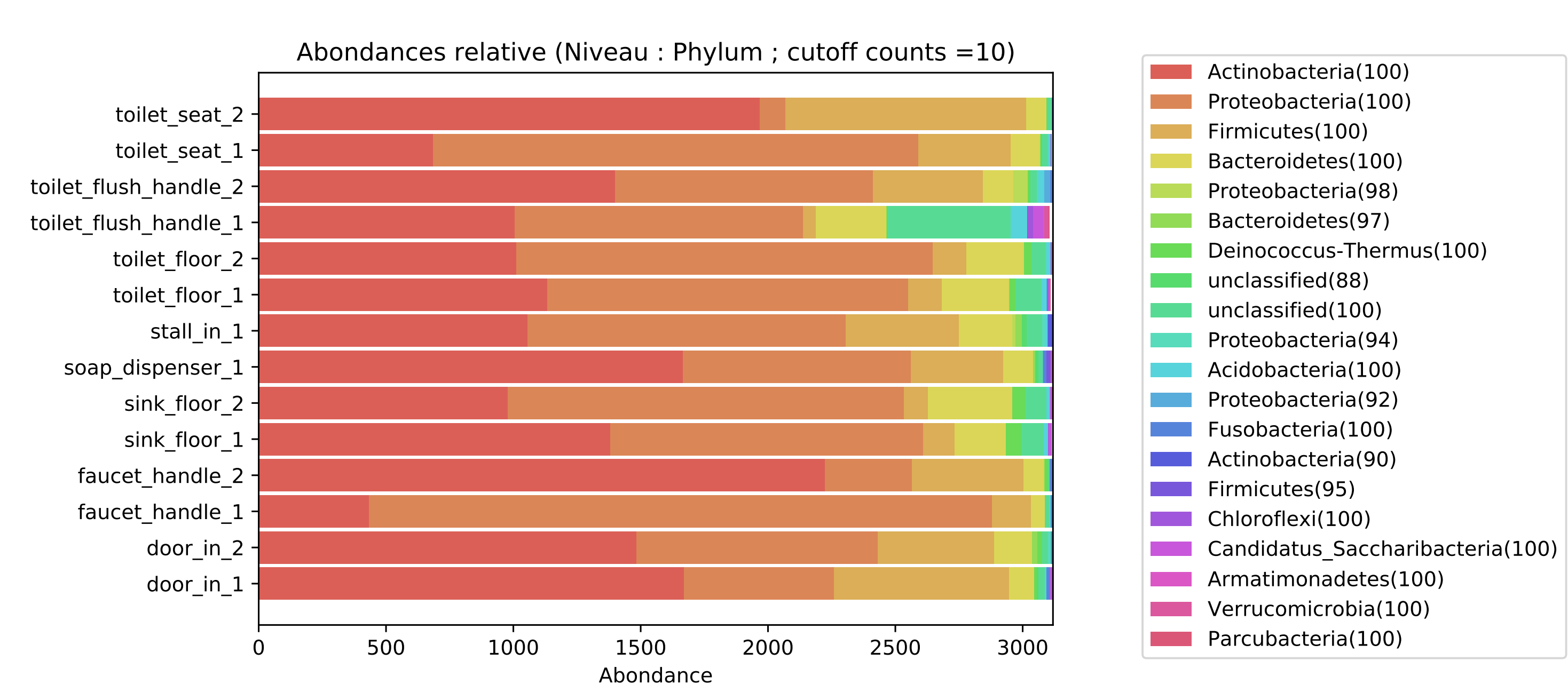
Marie Martelat et Johanna Galvis. M2 Bio-informatique, UCBL Lyon 1.

1. Introduction

En utilisant les données issus de la publication "Microbial Biogeography of Public Restroom Surfaces" de Flores *et al.*, nous voulons montrer les possibilités qu'offrent les packages Python pour la visualisation de données dans le contexte de la métagénomique. Nous avons à notre disposition une table de comptage de 14 échantillons. Les échantillons ont été collectés sur différentes surface de toilette. Surface touché par les mains (door_in, faucet_handle, toilet_flush_handle) par les pieds (sink_floor, toilet_floor) ou en contact direct avec la partie posterieur du corp(toilet_seat). Il peut être intéressant d'essayer de visualiser une différence de composition bactérienne dans ces différents type d'échantillons.

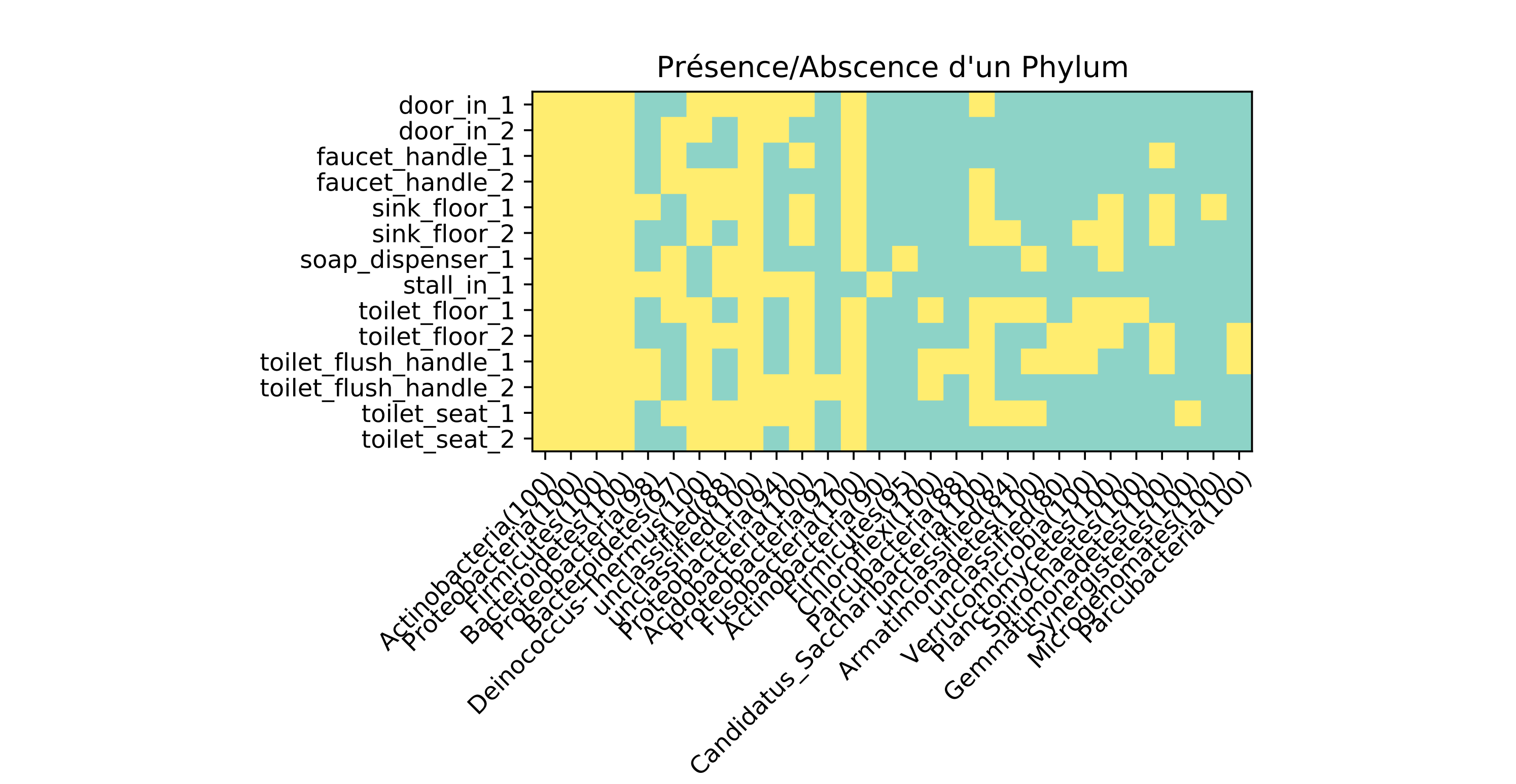
2.BarPlot

Abondance relative des phylums,sur différentes surfaces. Quatre taxons ressortent quelque soit la surface (**Actinobacteria**, **Proteobacteria**, **fimicutes**, **Bacteroidetes**).



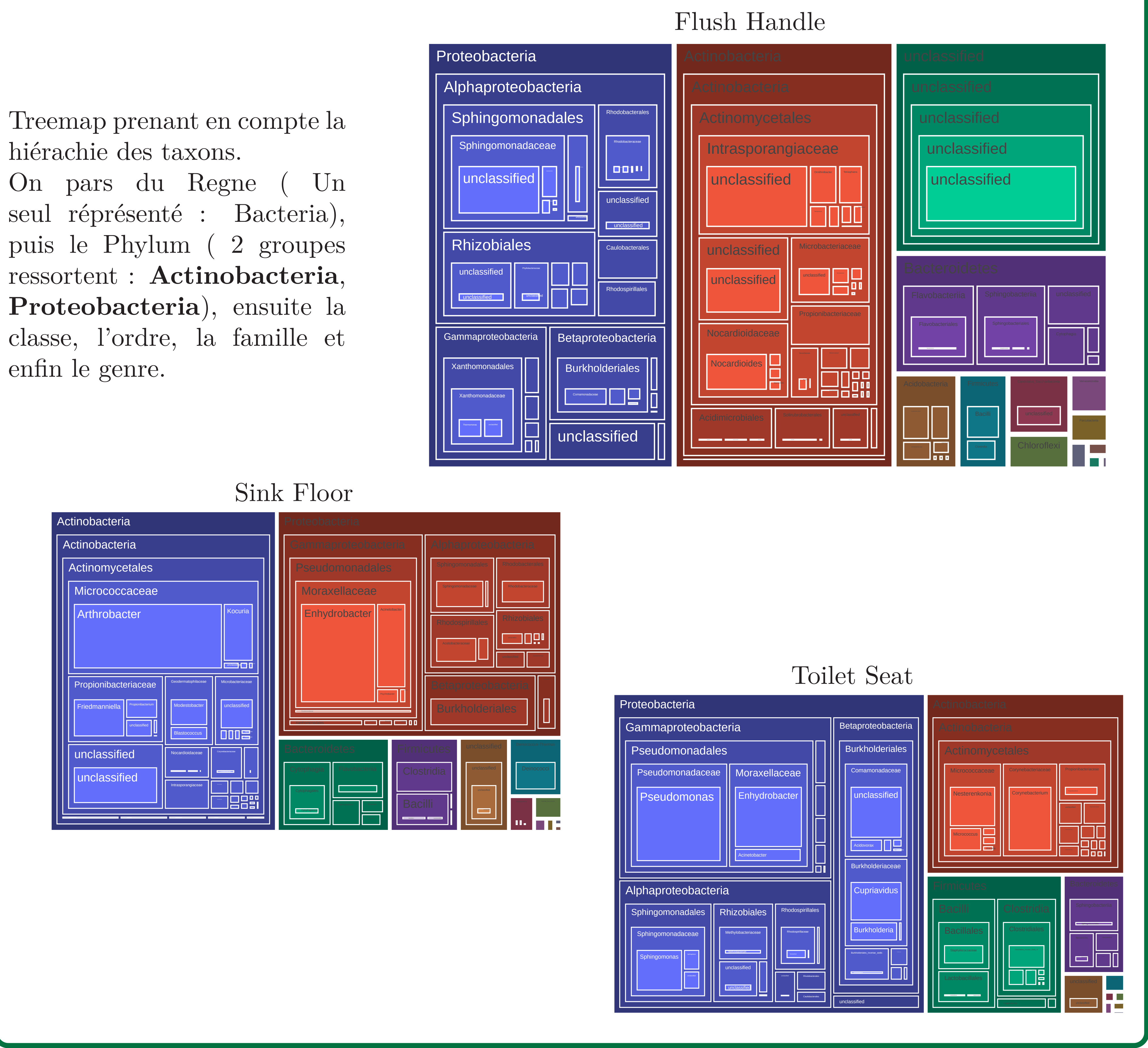
3.HeatMap

Visualisation de l'absence (ou présence) de taxon sur les différentes surfacess (Niveaux Phylum).



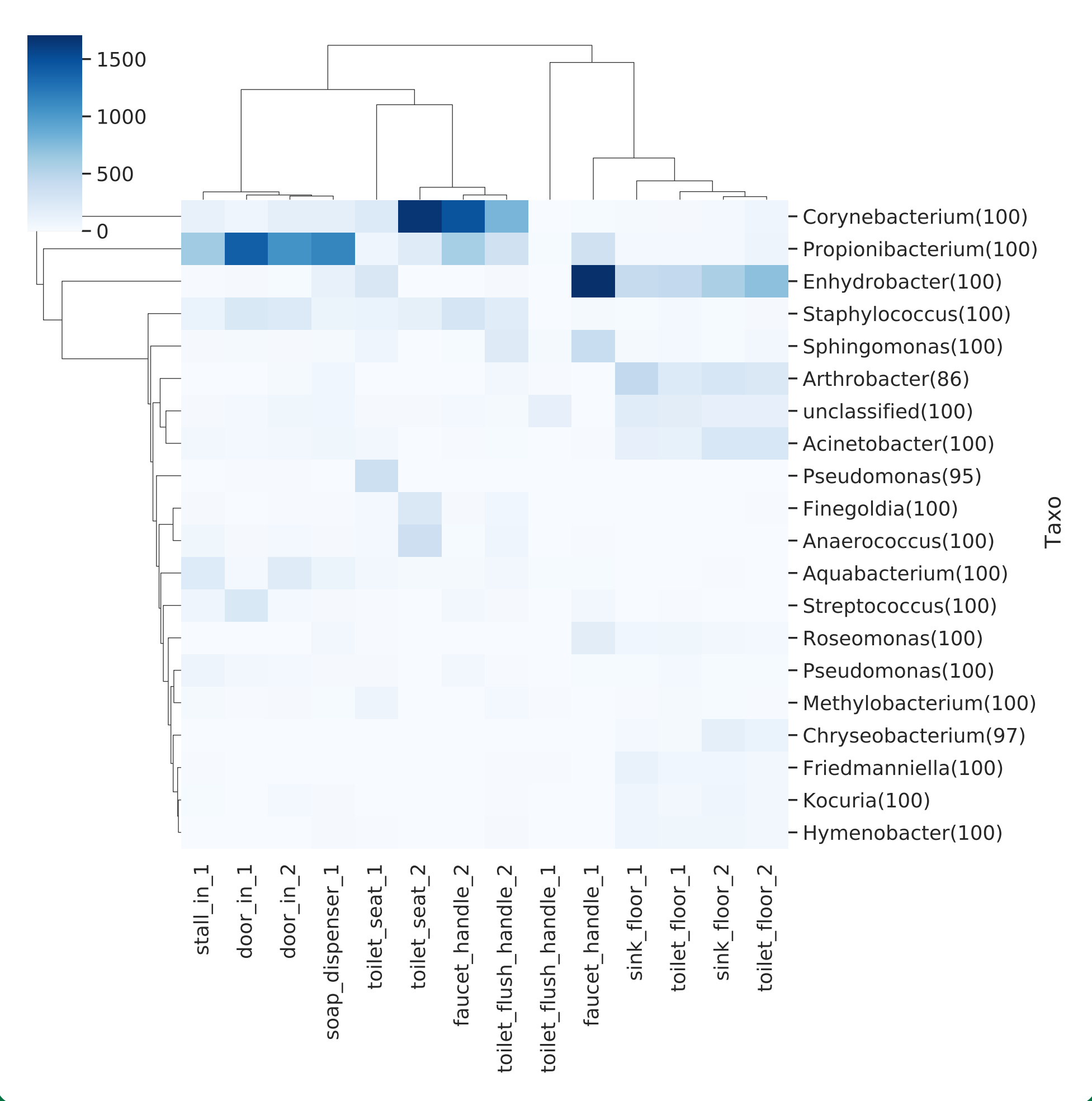
5.TreeMap

Treemap prenant en compte la hiérarchie des taxons. On pars du Regne (Un seul représenté : Bacteria), puis le Phylum (2 groupes ressortent : **Actinobacteria**, **Proteobacteria**), ensuite la classe, l'ordre, la famille et enfin le genre.



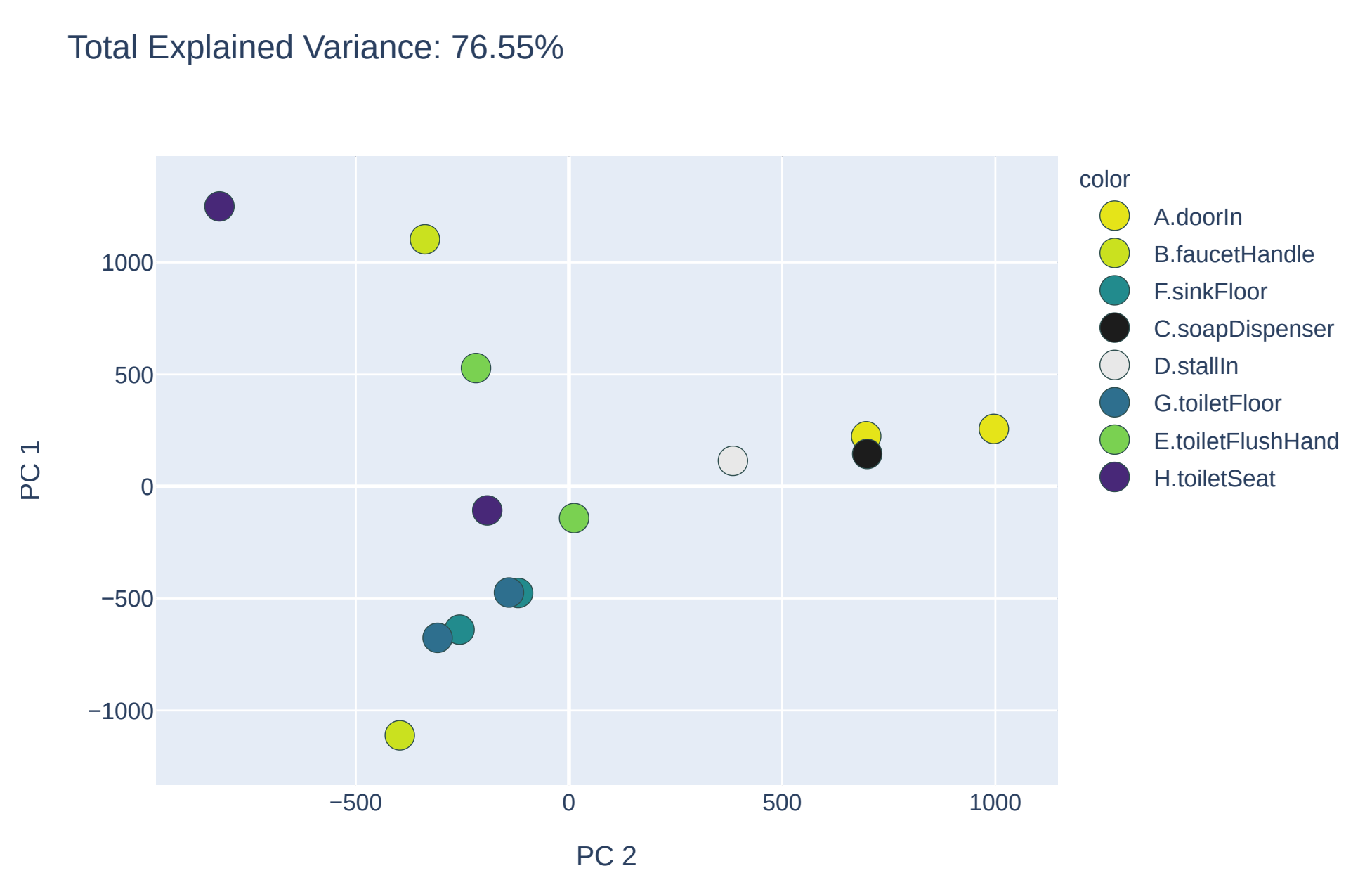
4. Clustering Hierarchique

Au niveau **genre**, le calcul des distances entre échantillons laisse apprécier des groupements selon la surface corporelle impliquée.



6. ACP

Notez la proximité inattendue entre échantillons du sols (bleue), **toilet_flush_handle**(vert) et robinet(citron).



Conclusion

Des puissants outils de visualisation permettent la découverte d'information dans les données. Malgré les limitations propres au librairies utilisées, des *patterns* intéressants ont pu être identifiés. Des **particularités** des échantillons doivent être tenues en compte (une composition inattendue).