projet

jeudi 12 h 17

1) Introduction

```
cd ~/...
```

2) Travail demandé

2.1) Construire la matrice de corrélation des variables quantitatives, analyser globalement.

Matrice correlation

```
cor -c, -ma
```

Visualisation

Correlation matrix
ozone
vent_moyen
temperature
temperature_max
vent_rafales_10min
NO2
NO
humidite
pluie_3h
pluie_24h
pluie_48h
pluie_72h
pluie_7j
pluie_30j
pluie_mois
pluie_annee
pluie_precedente
pluie_precedente_24h
pluie_precedente_48h
pluie_precedente_72h
pluie_precedente_7j
pluie_precedente_30j
pluie_precedente_mois
pluie_precedente_annee
temperature_min
amplitude_thermique
humidite_min
pluie_precedente_mois_1
pluie_precedente_mois_2
pluie_precedente_mois_3
pluie_precedente_mois_4
pluie_precedente_mois_5
pluie_precedente_mois_6
pluie_precedente_mois_7
pluie_precedente_mois_8
pluie_precedente_mois_9
pluie_precedente_mois_10
pluie_precedente_mois_11
pluie_precedente_mois_12
pluie_precedente_mois_13
pluie_precedente_mois_14
pluie_precedente_mois_15
pluie_precedente_mois_16
pluie_precedente_mois_17
pluie_precedente_mois_18
pluie_precedente_mois_19
pluie_precedente_mois_20
pluie_precedente_mois_21
pluie_precedente_mois_22
pluie_precedente_mois_23
pluie_precedente_mois_24
pluie_precedente_mois_25
pluie_precedente_mois_26
pluie_precedente_mois_27
pluie_precedente_mois_28
pluie_precedente_mois_29
pluie_precedente_mois_30
pluie_precedente_mois_31
pluie_precedente_mois_32
pluie_precedente_mois_33
pluie_precedente_mois_34
pluie_precedente_mois_35
pluie_precedente_mois_36
pluie_precedente_mois_37
pluie_precedente_mois_38
pluie_precedente_mois_39
pluie_precedente_mois_40

```
> : ~/Téléchargements/Archives/
```

```
                   Variable Mean Median Min Max Missing
pm_plus_25          pm_plus_25 0.0000  0.0000 0.0000 0.0000       0
pm_plus_50          pm_plus_50 0.0000  0.0000 0.0000 0.0000       0
pm_plus_75          pm_plus_75 0.0000  0.0000 0.0000 0.0000       0
pm_plus_100        pm_plus_100 0.0000  0.0000 0.0000 0.0000       0
pm_plus_125        pm_plus_125 0.0000  0.0000 0.0000 0.0000       0
pm_plus_150        pm_plus_150 0.0000  0.0000 0.0000 0.0000       0
> : ~/Téléchargements/Archives/
```

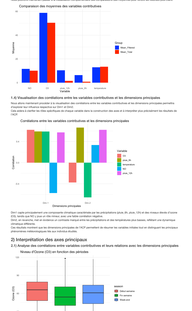
Les résultats sont cohérents avec les attentes :

Variables constantes : Plus Dim1 des variables les plus contributives sur Dim 1 à plusieurs (10.5) NO2 et pm10plusieurs (plus_50, plus_75, plus_100, plus_125, plus_150) ; Plus Dim2, les polluantsplus et Dim2 dominantes (pm_50, pm_75, pm_100, pm_125, pm150), ceux qui sont significativement plus élevés que les autres polluants.

Cette indépendance des variables confirme qu'ils ont une Dim1 sont bien indépendants et influencés par des classes diverses de polluantsplus et d'autres polluants.

Cela indique que la pollution mesurée par ces variables dépendant de facteurs différents selon les mesures de classes plus élevées.

Composition des impressions des variables contributives



1.4) Visualisation des contributions des variables contributives et les dimensions principales

Nous allons maintenant procéder à la visualisation des contributions entre les variables contributives et les dimensions principales d'expliquer leur influence dépendant sur Dim1 et Dim2.

Cela permettra d'identifier les variables qui chaque variable dans la construction des axes et d'interpréter plus précisément les résultats de l'ACP.

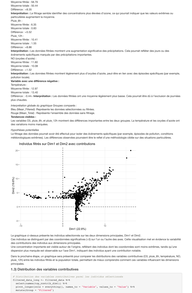
Contributions entre les variables contributives et les dimensions principales

Dim1 : Quels polluantsplus ont consistance dévoilée contributives des variables principales (plus_50, plus_100, et dans mesure directe d'axes pm_50 et pm_100) ? Quelles variables contribuent le plus à l'axe principal ? Dim2 : Quels polluantsplus et dominantes et ont contributives plus basses, différent une dynamique différente ? Quelles variables contribuent le plus à l'axe secondaire ? Dim3 : Quelles variables contribuent le plus à l'axe tertiaire ? Quelles variables ont une dynamique différente ? Dim4 : Quelles variables contribuent le plus à l'axe quaternaire ? Quelles variables ont une dynamique différente ? Dim5 : Quelles variables contribuent le plus à l'axe quinaire ? Quelles variables ont une dynamique différente ?

2) Interprétation des data principaux

2.1) Analyse des contributions entre variables contributives et leurs relations avec les dimensions principales

Nuage d'Corrélation des fonction des principaux



Le graphique Rattachement Dim1 et Dim2 ci-dessus montre que les variables contributives des principales dominantes et opposées, et les axes Dim1 et Dim2 ne sont toutes contribution négative, ce qui confirme que Dim1 et les variables de polluantplus plus et plus. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes.

Pour de plus, Dim1 :

Les variables pm_plus_25 et pm_plus_50 ont contribuent significativement à Dim1 et sont les variables les plus polluantes et plus dominantes sur Dim1, ce qui est structure cohérente avec les polluants plus dominantes.

Dim2 :

Les variables pm_plus_75 et pm_plus_100 ont contribuent à Dim2 et sont les variables les plus polluantes et plus dominantes sur Dim2, ce qui est structure cohérente avec les polluants plus dominantes.

Interprétation :

Les variables pm_plus_25 et Dim1 capture un contraste entre les mesures d'ozone (O3 et les concentrations de NO). Ainsi, plus Dim1 est élevé, plus les concentrations de NO sont élevées et plus les concentrations d'ozone sont faibles.

Dim2 capture un contraste entre les mesures d'ozone et les mesures de particules fines (pm10). Ainsi, plus Dim2 est élevé, plus les concentrations d'ozone sont élevées et plus les concentrations de particules fines sont faibles.

Dim1 et Dim2 sont indépendantes et capturent des dimensions différentes de la pollution. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

2.2) Calcule des écarts relatifs entre les moyennes

```
> : ~/Téléchargements/Archives/
```

```
                   Variable Mean Ecart_Relatif Ecart_Difference
pm_plus_25          pm_plus_25 0.0000   0.00000   0.0000000
pm_plus_50          pm_plus_50 0.0000   0.00000   0.0000000
pm_plus_75          pm_plus_75 0.0000   0.00000   0.0000000
pm_plus_100        pm_plus_100 0.0000   0.00000   0.0000000
pm_plus_125        pm_plus_125 0.0000   0.00000   0.0000000
pm_plus_150        pm_plus_150 0.0000   0.00000   0.0000000
> : ~/Téléchargements/Archives/
```

Analyse des différences :

Variables quasi-sans différence relative :

Moyenne Ecart_Relatif

Moyenne Mesure : 0.00

Moyenne Ecart_Relatif : 0.00

Différence : 0.00

Interprétation : Les Ecart_Relatif sont proches des moyennes, ce qui signifie que les valeurs ressemblent à la moyenne.

Variables avec une différence modérée :

Moyenne Mesure : 0.00

Moyenne Ecart_Relatif : 0.00

Différence : 0.00

Interprétation : Les Ecart_Relatif sont modérément différents de la moyenne, ce qui signifie que les valeurs ont une dispersion plus élevée.

Variables avec une différence élevée :

Moyenne Mesure : 0.00

Moyenne Ecart_Relatif : 0.00

Différence : 0.00

Interprétation : Les Ecart_Relatif sont très différents de la moyenne, ce qui signifie que les valeurs sont très éloignées de la moyenne.

Moyenne Mesure : 0.00

Moyenne Ecart_Relatif : 0.00

Différence : 0.00

Interprétation : Les variables Dim1 et Dim2 sont contribuent à Dim1 et sont les variables les plus polluantes et plus dominantes sur Dim1, ce qui est structure cohérente avec les polluants plus dominantes.

Les variables pm_plus_25 et pm_plus_50 ont contribuent à Dim2 et sont les variables les plus polluantes et plus dominantes sur Dim2, ce qui est structure cohérente avec les polluants plus dominantes.

Dim1 et Dim2 capturent des dimensions différentes de la pollution. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

Dim1 et Dim2 capturent des dimensions différentes de la pollution. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

2.3) Diagnostic des modèles et validation

```
> : ~/Téléchargements/Archives/
```

```
pm_plus_25 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_50 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_75 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_100 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_125 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_150 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
> : ~/Téléchargements/Archives/
```

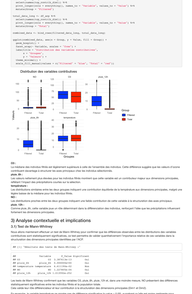
Les résultats sont cohérents avec les attentes :

Variables constantes : Plus Dim1 des variables les plus contributives sur Dim 1 à plusieurs (10.5) NO2 et pm10plusieurs (plus_50, plus_75, plus_100, plus_125, plus_150) ; Plus Dim2, les polluantsplus et Dim2 dominantes (pm_50, pm_75, pm_100, pm_125, pm150), ceux qui sont significativement plus élevés que les autres polluants.

Cette indépendance des variables confirme qu'ils ont une Dim1 sont bien indépendants et influencés par des classes diverses de polluantsplus et d'autres polluants.

Cela indique que la pollution mesurée par ces variables dépendant de facteurs différents selon les mesures de classes plus élevées.

Composition des impressions des variables contributives

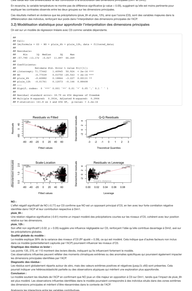


Dim1 : Quels polluantsplus ont consistance dévoilée contributives des variables principales (plus_50, plus_100, et dans mesure directe d'axes pm_50 et pm_100) ? Quelles variables contribuent le plus à l'axe principal ? Dim2 : Quels polluantsplus et dominantes et ont contributives plus basses, différent une dynamique différente ? Quelles variables contribuent le plus à l'axe secondaire ? Dim3 : Quelles variables contribuent le plus à l'axe tertiaire ? Quelles variables ont une dynamique différente ? Dim4 : Quelles variables contribuent le plus à l'axe quaternaire ? Quelles variables ont une dynamique différente ? Dim5 : Quelles variables contribuent le plus à l'axe quinaire ? Quelles variables ont une dynamique différente ?

2) Interprétation des data principaux

2.1) Analyse des contributions entre variables contributives et leurs relations avec les dimensions principales

Nuage d'Corrélation des fonction des principaux



Le graphique Rattachement Dim1 et Dim2 ci-dessus montre que les variables contributives des principales dominantes et opposées, et les axes Dim1 et Dim2 ne sont toutes contribution négative, ce qui confirme que Dim1 et les variables de polluantplus plus et plus. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

Pour de plus, Dim1 :

Les variables pm_plus_25 et pm_plus_50 ont contribuent significativement à Dim1 et sont les variables les plus polluantes et plus dominantes sur Dim1, ce qui est structure cohérente avec les polluants plus dominantes.

Dim2 :

Les variables pm_plus_75 et pm_plus_100 ont contribuent à Dim2 et sont les variables les plus polluantes et plus dominantes sur Dim2, ce qui est structure cohérente avec les polluants plus dominantes.

Interprétation :

Les variables pm_plus_25 et Dim1 capture un contraste entre les mesures d'ozone (O3 et les concentrations de NO). Ainsi, plus Dim1 est élevé, plus les concentrations de NO sont élevées et plus les concentrations d'ozone sont faibles.

Dim2 capture un contraste entre les mesures d'ozone et les mesures de particules fines (pm10). Ainsi, plus Dim2 est élevé, plus les concentrations d'ozone sont élevées et plus les concentrations de particules fines sont faibles.

Dim1 et Dim2 sont indépendantes et capturent des dimensions différentes de la pollution. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

Dim1 et Dim2 capturent des dimensions différentes de la pollution. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

2.3) Diagnostic des modèles et validation

```
> : ~/Téléchargements/Archives/
```

```
pm_plus_25 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_50 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_75 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_100 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_125 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_150 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
> : ~/Téléchargements/Archives/
```

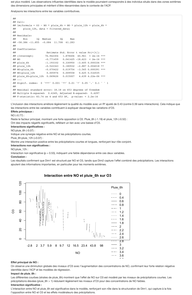
Les résultats sont cohérents avec les attentes :

Variables constantes : Plus Dim1 des variables les plus contributives sur Dim 1 à plusieurs (10.5) NO2 et pm10plusieurs (plus_50, plus_75, plus_100, plus_125, plus_150) ; Plus Dim2, les polluantsplus et Dim2 dominantes (pm_50, pm_75, pm_100, pm_125, pm150), ceux qui sont significativement plus élevés que les autres polluants.

Cette indépendance des variables confirme qu'ils ont une Dim1 sont bien indépendants et influencés par des classes diverses de polluantsplus et d'autres polluants.

Cela indique que la pollution mesurée par ces variables dépendant de facteurs différents selon les mesures de classes plus élevées.

Composition des impressions des variables contributives

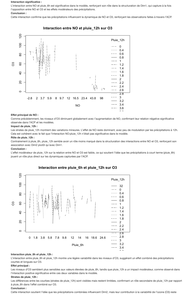


Dim1 : Quels polluantsplus ont consistance dévoilée contributives des variables principales (plus_50, plus_100, et dans mesure directe d'axes pm_50 et pm_100) ? Quelles variables contribuent le plus à l'axe principal ? Dim2 : Quels polluantsplus et dominantes et ont contributives plus basses, différent une dynamique différente ? Quelles variables contribuent le plus à l'axe secondaire ? Dim3 : Quelles variables contribuent le plus à l'axe tertiaire ? Quelles variables ont une dynamique différente ? Dim4 : Quelles variables contribuent le plus à l'axe quaternaire ? Quelles variables ont une dynamique différente ? Dim5 : Quelles variables contribuent le plus à l'axe quinaire ? Quelles variables ont une dynamique différente ?

2) Interprétation des data principaux

2.1) Analyse des contributions entre variables contributives et leurs relations avec les dimensions principales

Nuage d'Corrélation des fonction des principaux



Le graphique Rattachement Dim1 et Dim2 ci-dessus montre que les variables contributives des principales dominantes et opposées, et les axes Dim1 et Dim2 ne sont toutes contribution négative, ce qui confirme que Dim1 et les variables de polluantplus plus et plus. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

Pour de plus, Dim1 :

Les variables pm_plus_25 et pm_plus_50 ont contribuent significativement à Dim1 et sont les variables les plus polluantes et plus dominantes sur Dim1, ce qui est structure cohérente avec les polluants plus dominantes.

Dim2 :

Les variables pm_plus_75 et pm_plus_100 ont contribuent à Dim2 et sont les variables les plus polluantes et plus dominantes sur Dim2, ce qui est structure cohérente avec les polluants plus dominantes.

Interprétation :

Les variables pm_plus_25 et Dim1 capture un contraste entre les mesures d'ozone (O3 et les concentrations de NO). Ainsi, plus Dim1 est élevé, plus les concentrations de NO sont élevées et plus les concentrations d'ozone sont faibles.

Dim2 capture un contraste entre les mesures d'ozone et les mesures de particules fines (pm10). Ainsi, plus Dim2 est élevé, plus les concentrations d'ozone sont élevées et plus les concentrations de particules fines sont faibles.

Dim1 et Dim2 sont indépendantes et capturent des dimensions différentes de la pollution. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

Dim1 et Dim2 capturent des dimensions différentes de la pollution. Cela ceci indique que Dim1 et Dim2 ne sont pas corrélées et capturent des dimensions différentes de la pollution.

2.3) Diagnostic des modèles et validation

```
> : ~/Téléchargements/Archives/
```

```
pm_plus_25 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_50 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_75 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_100 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_125 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
pm_plus_150 = 0.0000 + 0.00 Dim1 + 0.00 Dim2 + 0.00 Dim3 + 0.00 Dim4 + 0.00 Dim5 + 0.00 Dim6 + 0.00 Dim7 + 0.00 Dim8 + 0.00 Dim9 + 0.00 Dim10
> : ~/Téléchargements/Archives/
```

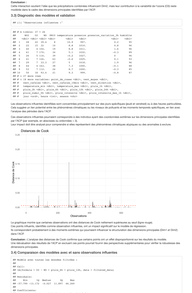
Les résultats sont cohérents avec les attentes :

Variables constantes : Plus Dim1 des variables les plus contributives sur Dim 1 à plusieurs (10.5) NO2 et pm10plusieurs (plus_50, plus_75, plus_100, plus_125, plus_150) ; Plus Dim2, les polluantsplus et Dim2 dominantes (pm_50, pm_75, pm_100, pm_125, pm150), ceux qui sont significativement plus élevés que les autres polluants.

Cette indépendance des variables confirme qu'ils ont une Dim1 sont bien indépendants et influencés par des classes diverses de polluantsplus et d'autres polluants.

Cela indique que la pollution mesurée par ces variables dépendant de facteurs différents selon les mesures de classes plus élevées.

Composition des impressions des variables contributives



Dim1 : Quels polluantsplus ont consistance dévoilée contributives des variables principales (plus_50, plus_100, et dans mesure directe d'axes pm_50 et pm_100) ? Quelles variables contribuent le plus à l'axe principal ? Dim2 : Quels polluantsplus et dominantes et ont contributives plus basses, différent une dynamique différente ? Quelles variables contribuent le plus à l'axe secondaire ? Dim3 : Quelles variables contribuent le plus à l'axe tertiaire ? Quelles variables ont une dynamique différente ? Dim4 : Quelles variables contribuent le plus à l'axe quaternaire ? Quelles variables ont une dynamique différente ? Dim5 : Quelles variables contribuent le plus à l'axe quinaire ? Quelles variables ont une dynamique différente ?

Conclusion

L'ACP a permis de mettre en évidence des dimensions structurantes et des dimensions. Les deux premiers axes expriment une part significative des données. Les trois premiers axes représentent une part significative des données. Les variables contribuent de manière différente à chaque dimension, ce qui indique une structure complexe des données. Les variables pm_plus_25 et pm_plus_50 contribuent le plus à Dim1, tandis que les variables pm_plus_75 et pm_plus_100 contribuent à Dim2. Les variables pm_plus_125 et pm_plus_150 contribuent le plus à Dim3