EDA for Salicylates and Benzophenones

Daihao(Will) Wu

2024-07-08

```
## [1] "Hexyl salicylate"        "2-ethylhexyl salicylate"
## [3] "Benzyl salicylate"       "\nBP"
## [5] "BP-3"                    "BP-8"
```

Above 6 chemicals have detection frequency above 50%.

Three of them are from Salicylates family. All three chemicals from BP family have detection frequency above 50%.

There are in total 44 houses in our dataset The following are the houses with 0 to 12 months period recorded:

"NHAQS-028" "NHAQS-029" "NHAQS-030" "NHAQS-031" "NHAQS-032" "NHAQS-034" "NHAQS-036"

There are 8 houses have 1 period recorded.

10 houses have 2 periods recorded.

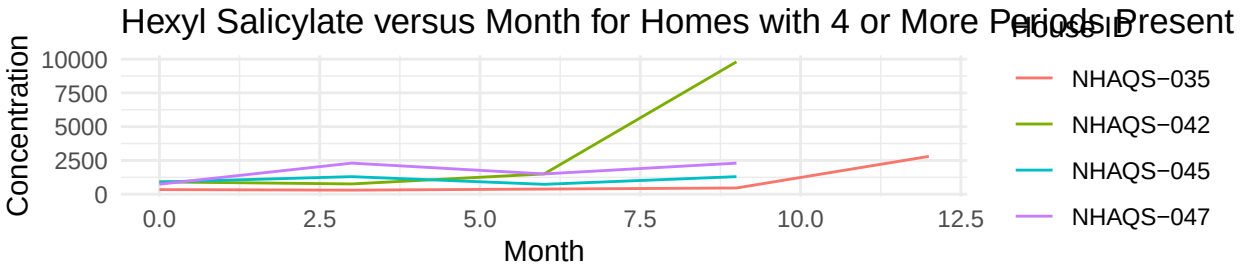
7 houses have 3 periods recorded.

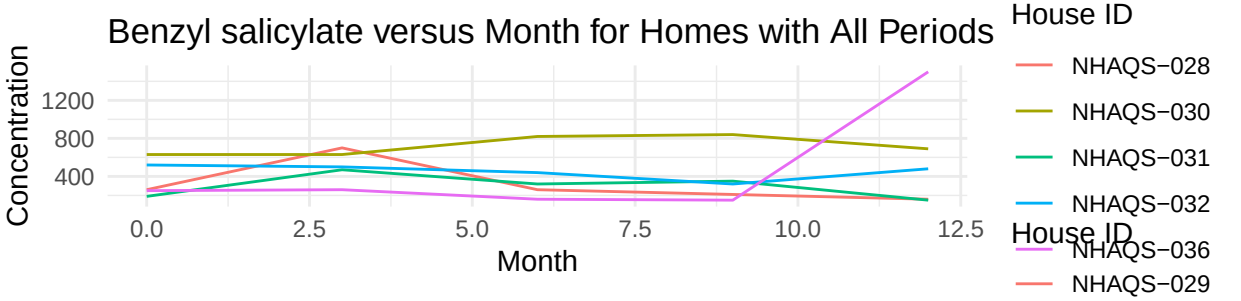
12 houses have 4 periods recorded.

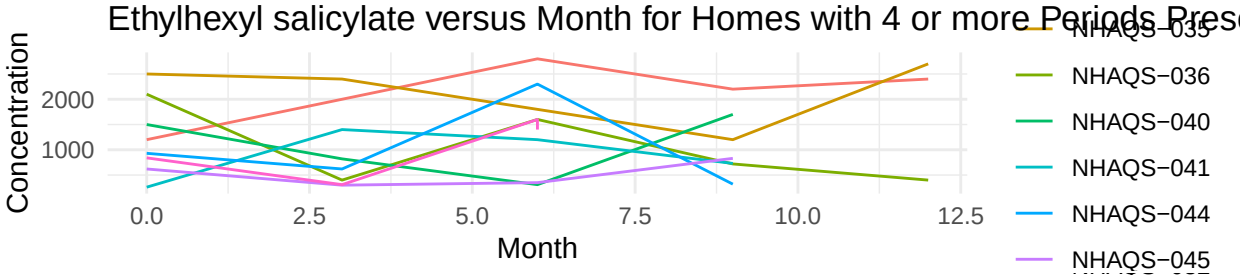
5 houses have 5 periods (0, 3, 6, 9, 12) recorded.

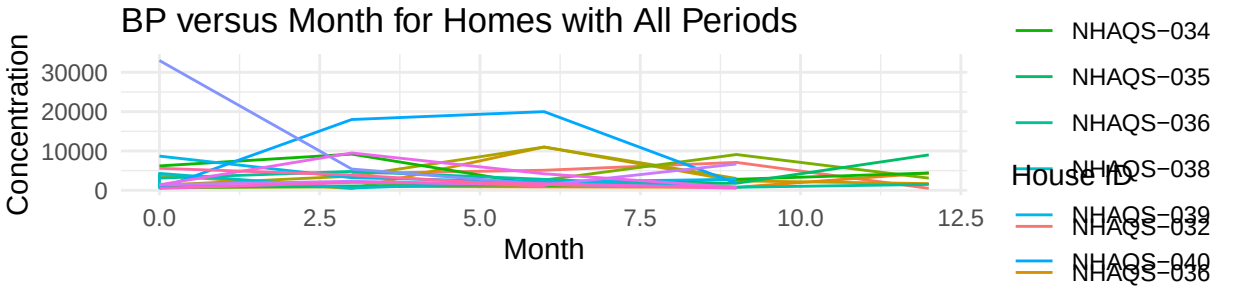
2 houses have 6 periods (0, 3, 6, 9, 12, other unusual period value e.g. Blank) recorded.

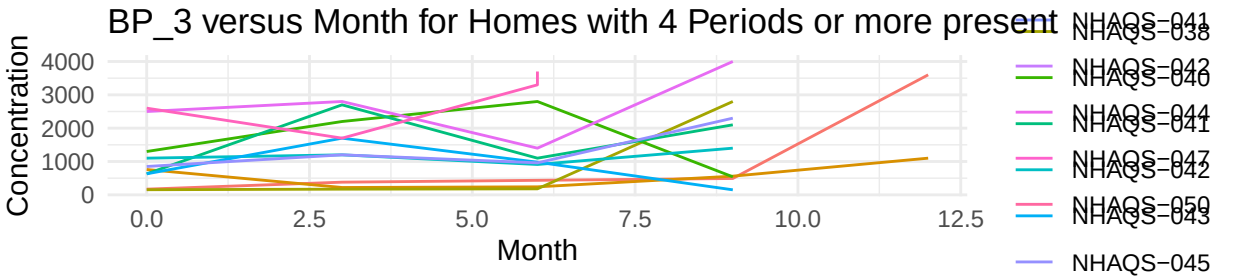
Note that House "NHAQS-38" has two rows of data recorded for the same period ('0') with the same Sample_ID. This might be a data collection/recording error.

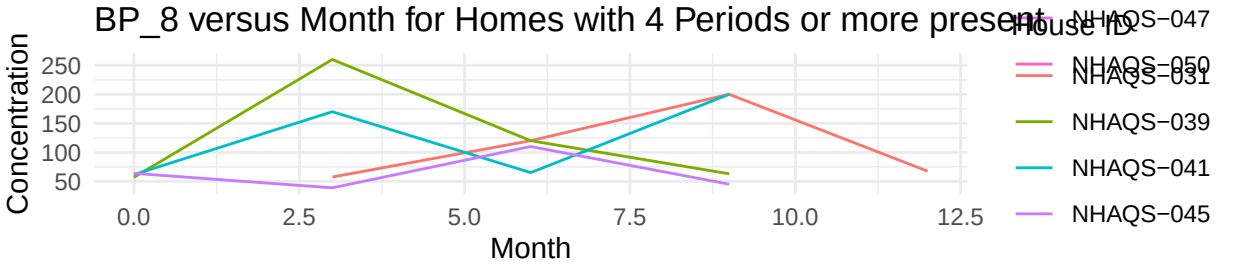

Hexyl Salicylate versus Month for Homes with 4 or More Periods Present
Concentration
10000
7500
5000
2500
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
House ID
NHAQS−035
NHAQS−042
NHAQS−045
NHAQS−047


Benzyl salicylate versus Month for Homes with All Periods
Concentration
1200
800
400
0.0
2.5
5.0
7.5
10.0
12.5
Month
House ID
NHAQS−028
NHAQS−030
NHAQS−031
NHAQS−032
NHAQS−036
NHAQS−029


Ethylhexyl salicylate versus Month for Homes with 4 or more Periods Pres
Concentration
2000
1000
0.0
2.5
5.0
7.5
10.0
12.5
Month
NHAQS−035
NHAQS−036
NHAQS−040
NHAQS−041
NHAQS−044
NHAQS−045


BP versus Month for Homes with All Periods
Concentration
30000
20000
10000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
NHAQS−034
NHAQS−035
NHAQS−036
NHAQS−038
House ID
NHAQS−039
NHAQS−032
NHAQS−040
NHAQS−030
NHAQS−041
NHAQS−038
NHAQS−042
NHAQS−040
NHAQS−044
NHAQS−041
NHAQS−047
NHAQS−042
NHAQS−050
NHAQS−043
NHAQS−045


BP_3 versus Month for Homes with 4 Periods or more present
Concentration
4000
3000
2000
1000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
NHAQS−047
House ID
NHAQS−050
NHAQS−031
NHAQS−039
NHAQS−041
NHAQS−045


BP_8 versus Month for Homes with 4 Periods or more present
Concentration
250
200
150
100
50
0.0
2.5
5.0
7.5
10.0
12.5
Month

From the above line plots, we can notice that:

(1) Increasing trend for Hexyl salicylate after the period of 3 for most houses

(2) The concentration level for BP fluctuate by a lot across periods

(3) There are many outliers for EthylHexyl salicylate (some go up to 5 times of the average concentration level)

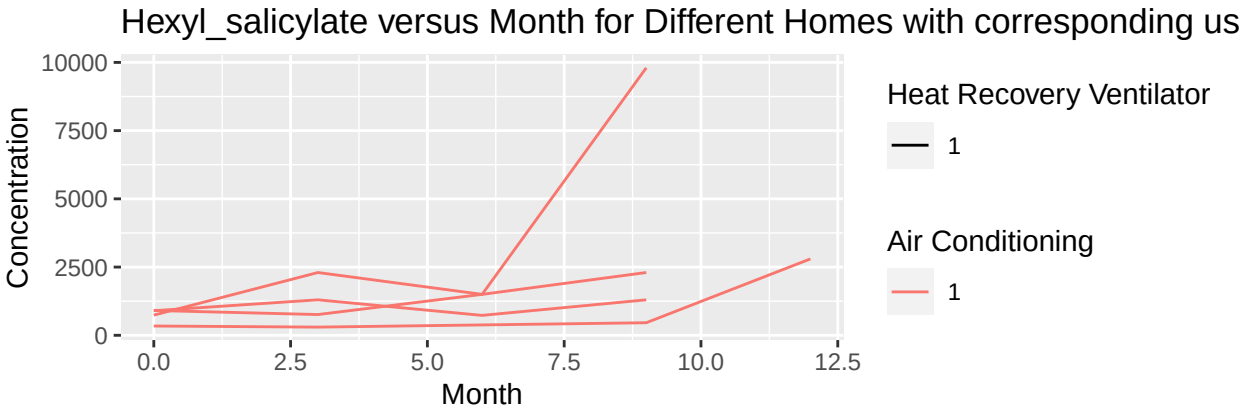
(4) For BP_3, if we look at the houses with 4 or more periods present, there seems to have a slight increase trend over time. However, mostly on the last recorded period.

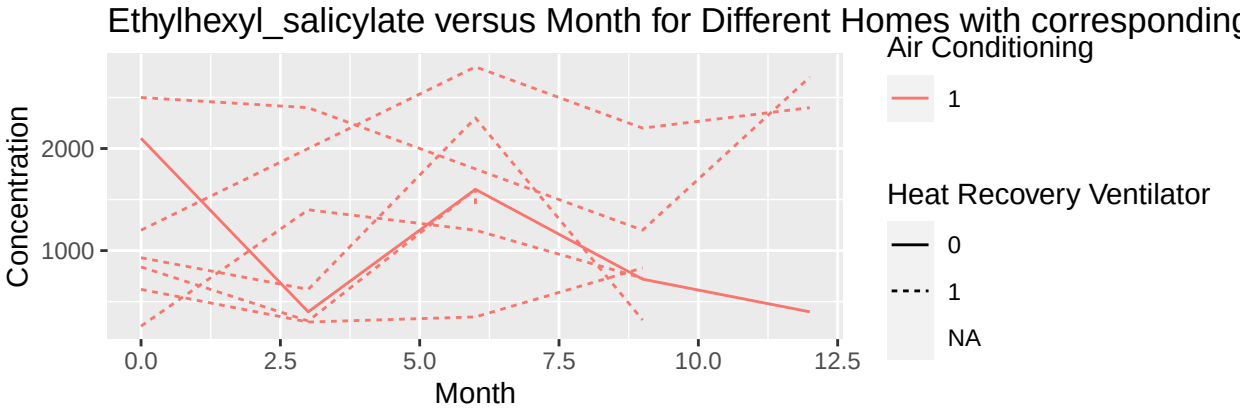
(5) For BP_8, there seems to have a slight decrease trend over time for most houses

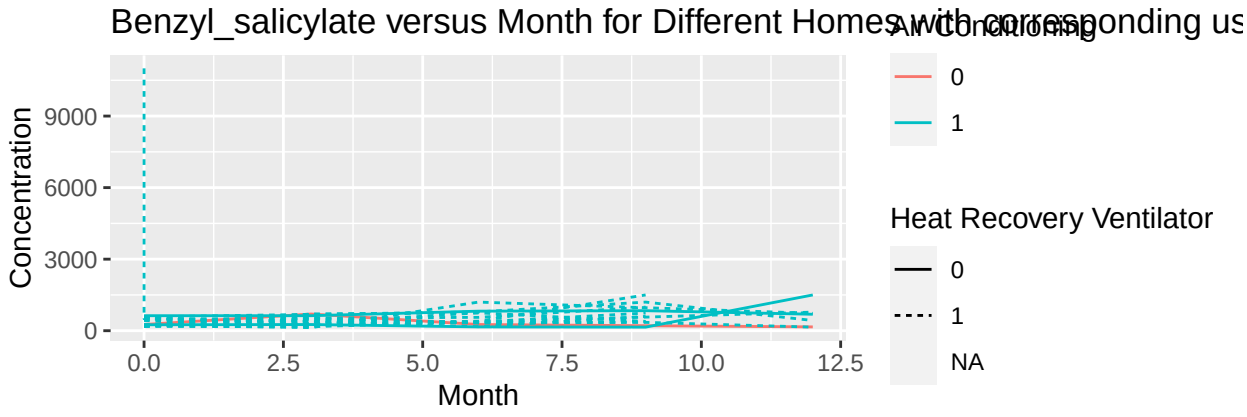
Jul 13, 2024

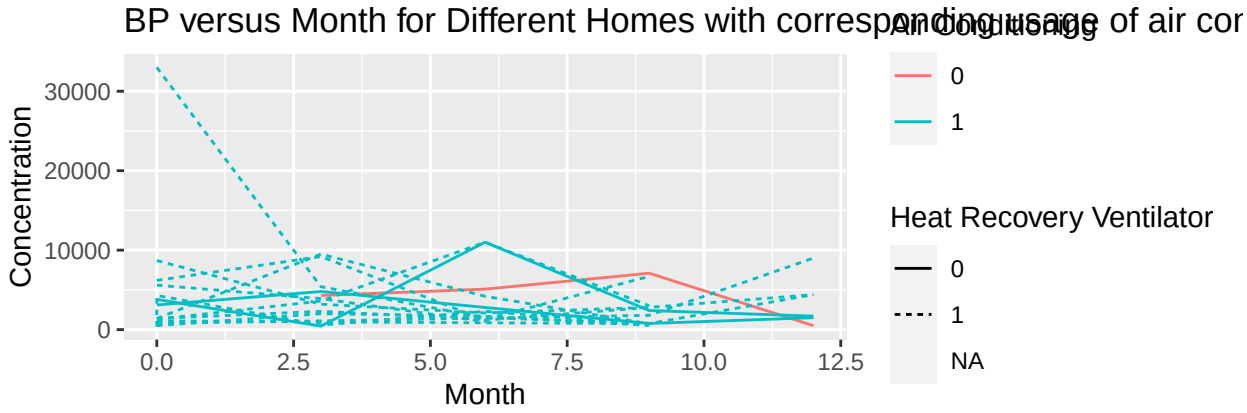
Look at factors that may affect levels of concentrations and do the line plots (as in the report)

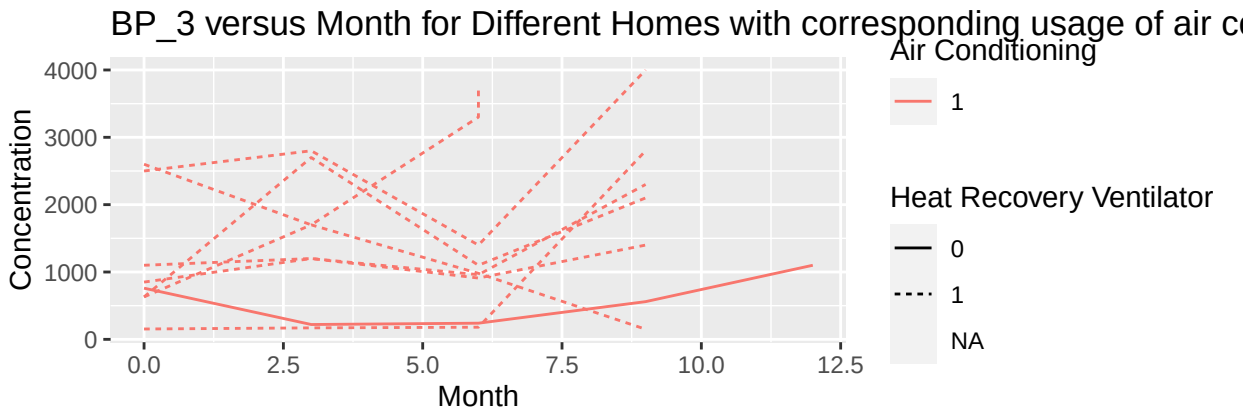
Cross-correlation between the chemicals.

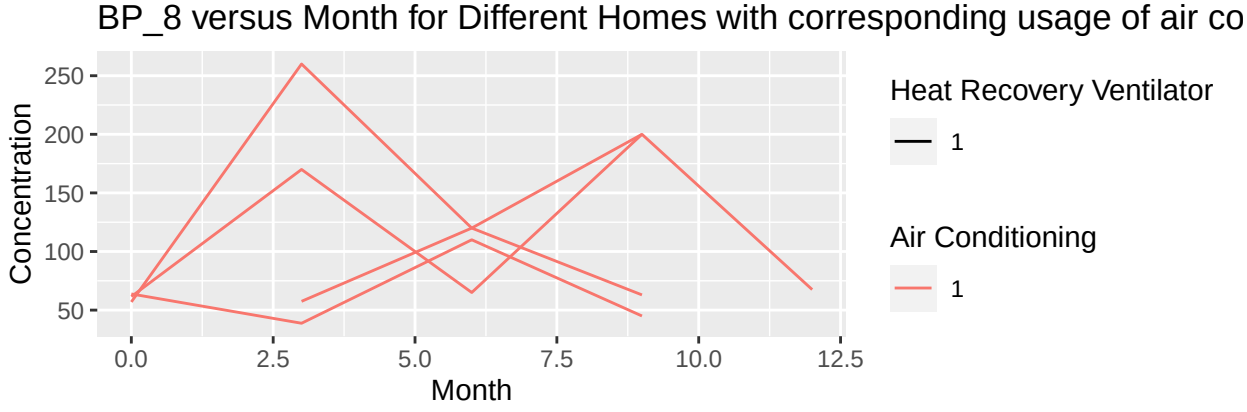


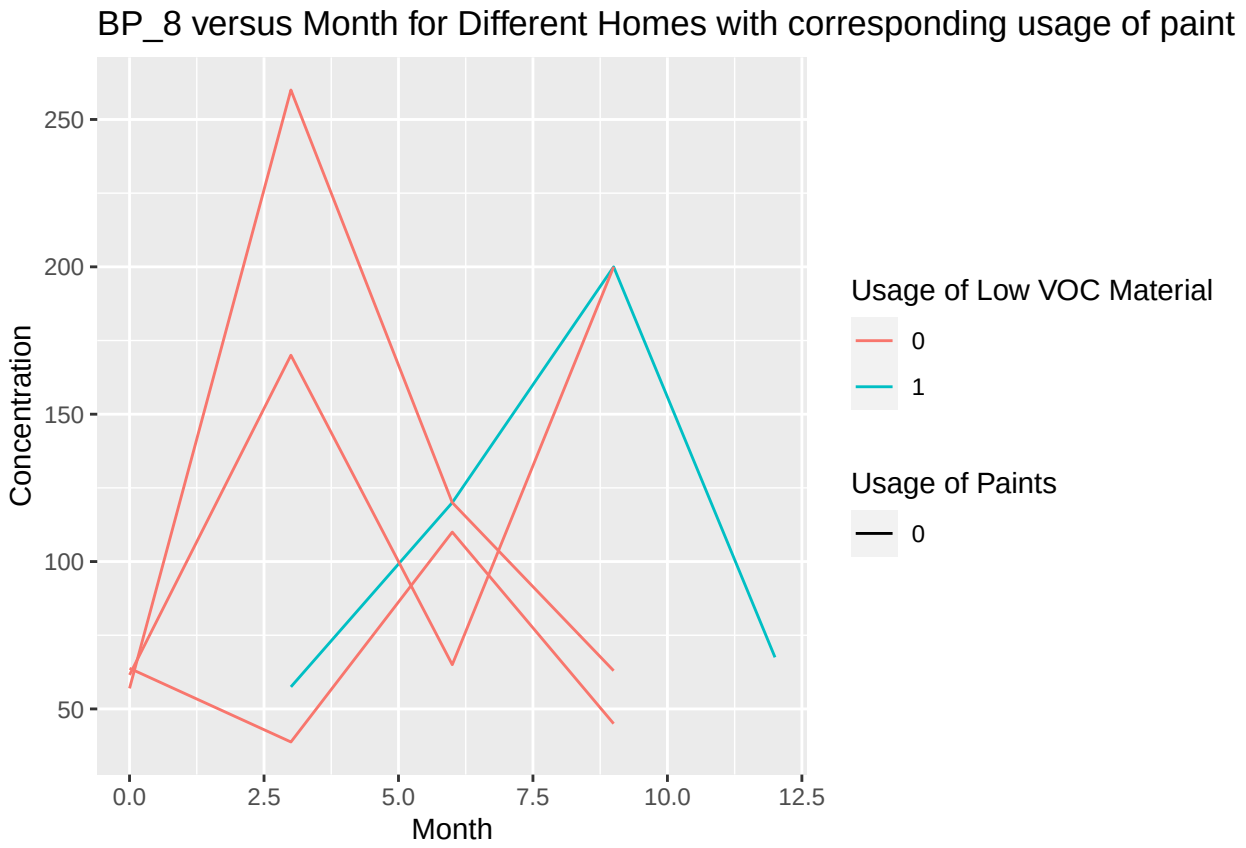


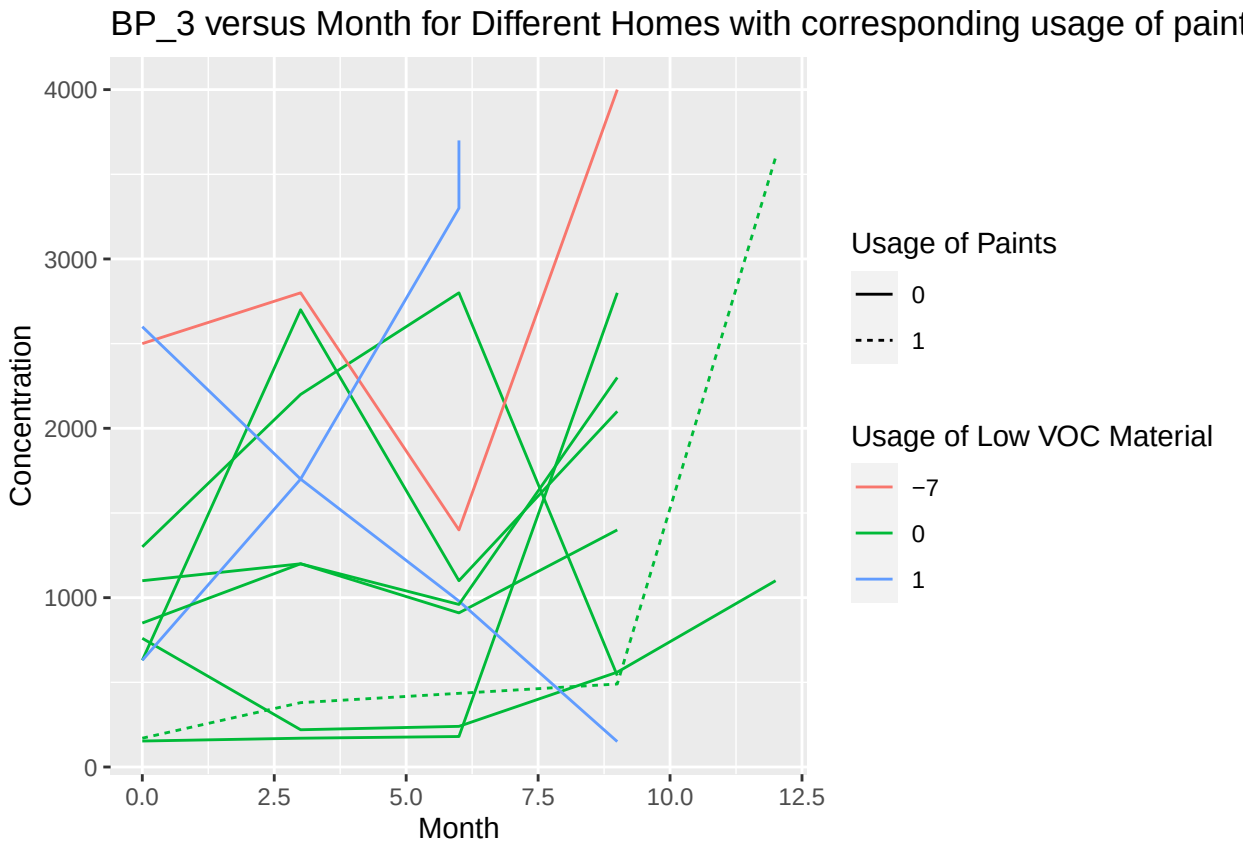

Benzyl_salicylate versus Month for Different Homes with corresponding us
Concentration
9000
6000
3000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Air Conditioning
0
1
Heat Recovery Ventilator
0
1
NA

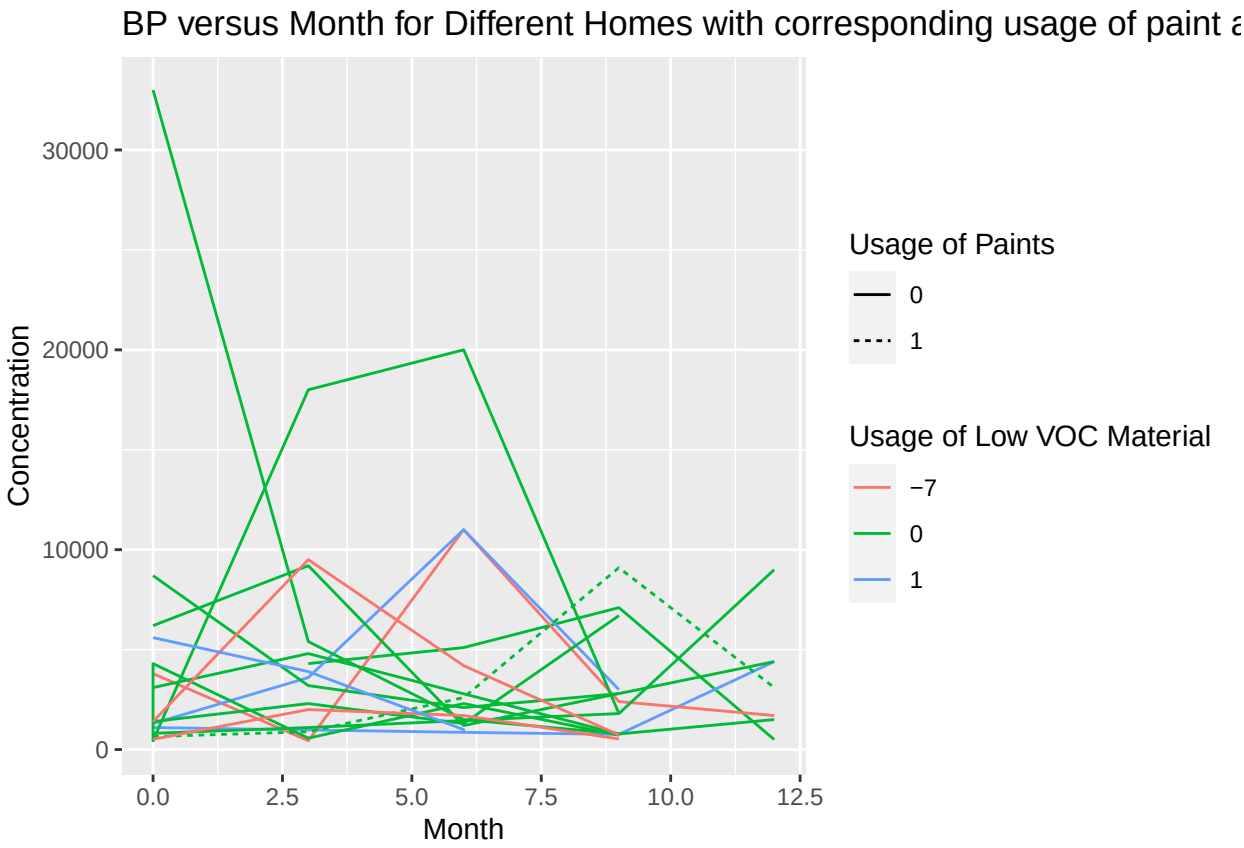

BP versus Month for Different Homes with corresponding usage of air cor
Concentration
30000
20000
10000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Air Conditioning
0
1
Heat Recovery Ventilator
0
1
NA

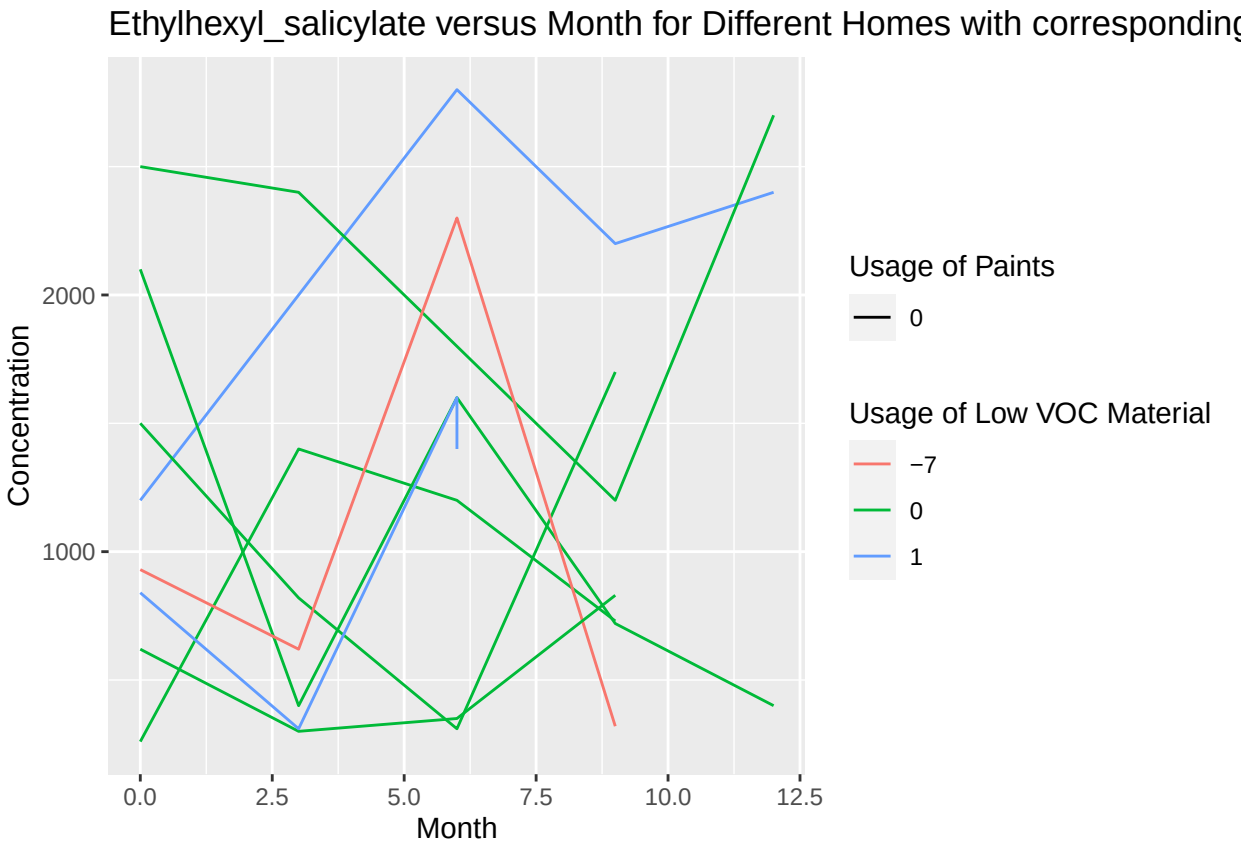

BP_3 versus Month for Different Homes with corresponding usage of air c
Concentration
4000
3000
2000
1000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Air Conditioning
1
Heat Recovery Ventilator
0
1
NA

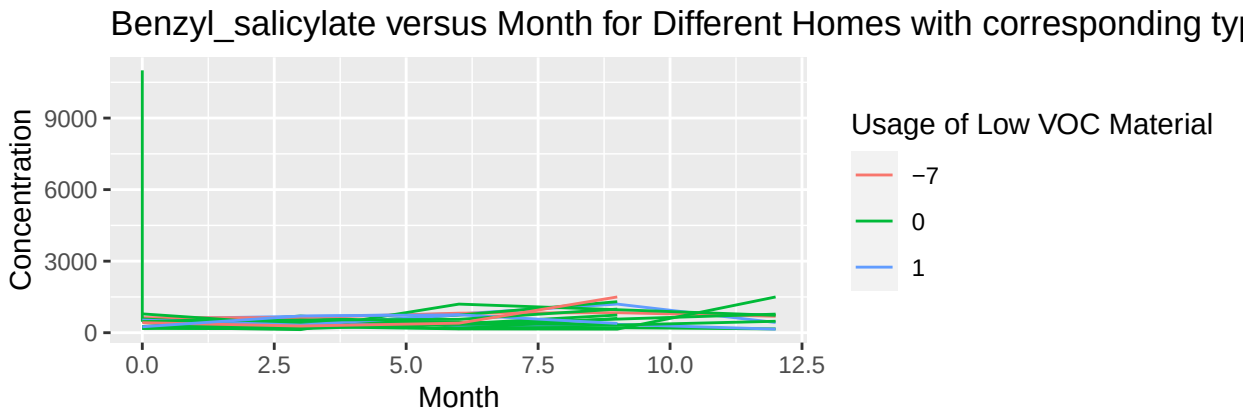

BP_8 versus Month for Different Homes with corresponding usage of air co
Concentration
250
200
150
100
50
0.0
2.5
5.0
7.5
10.0
12.5
Month
Heat Recovery Ventilator
1
Air Conditioning
1

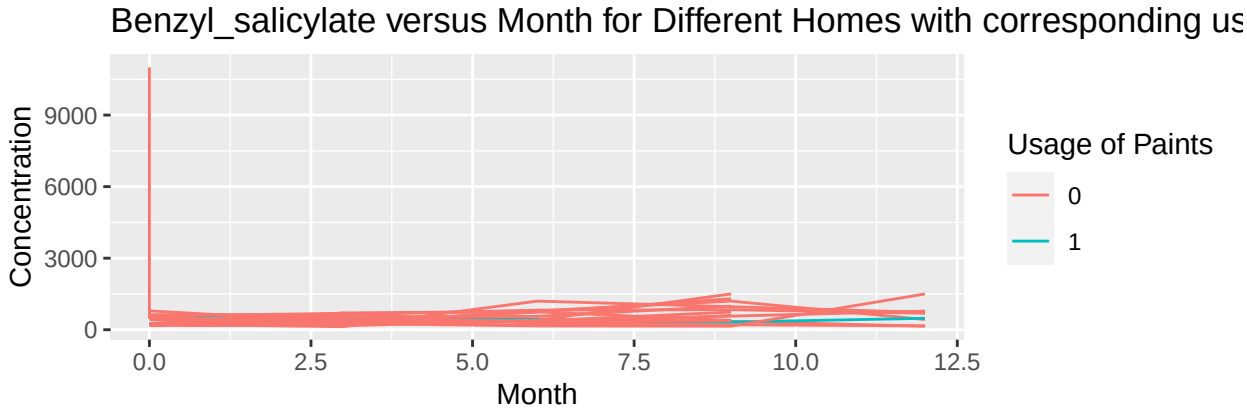

BP_8 versus Month for Different Homes with corresponding usage of paint
Concentration
250
200
150
100
50
0.0
2.5
5.0
7.5
10.0
12.5
Month
Usage of Low VOC Material
0
1
Usage of Paints
0

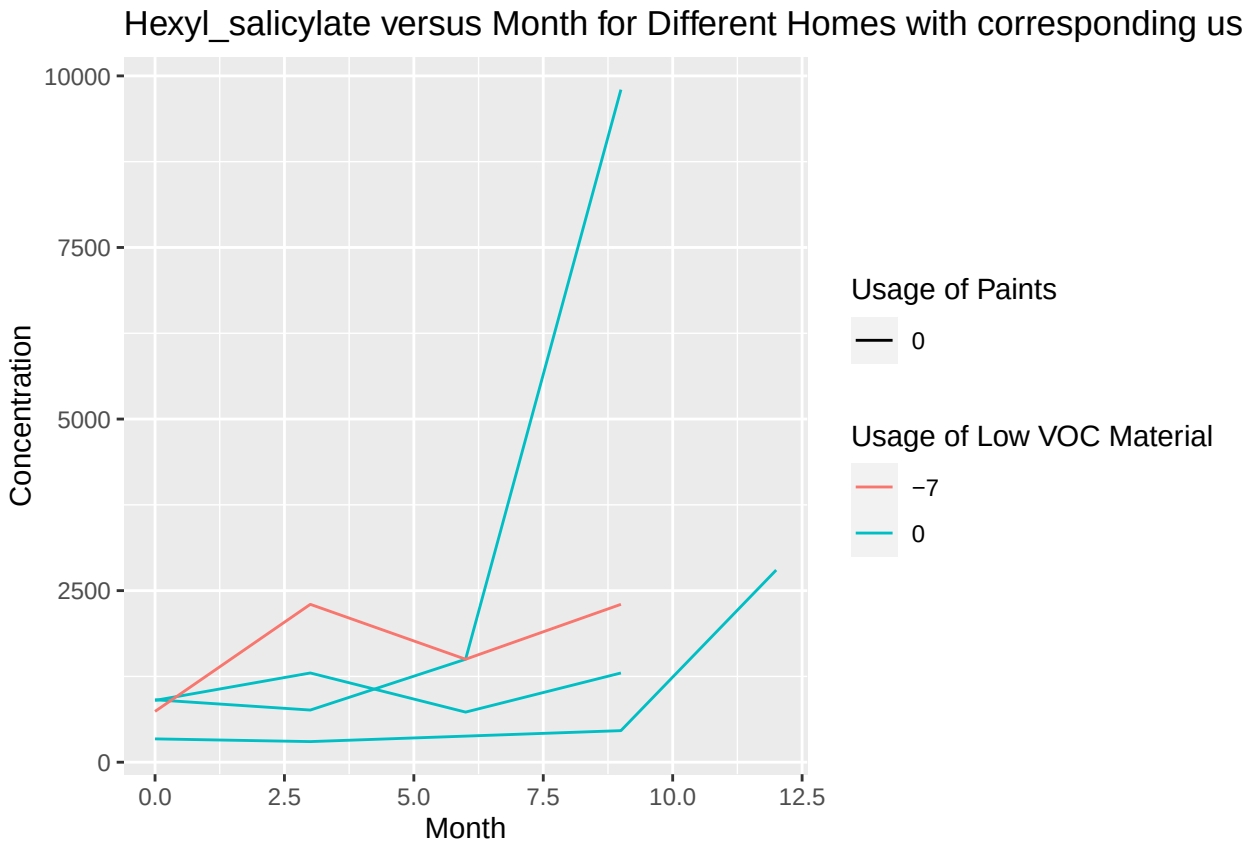

BP_3 versus Month for Different Homes with corresponding usage of paint
Concentration
4000
3000
2000
1000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Usage of Paints
0
1
Usage of Low VOC Material
−7
0
1

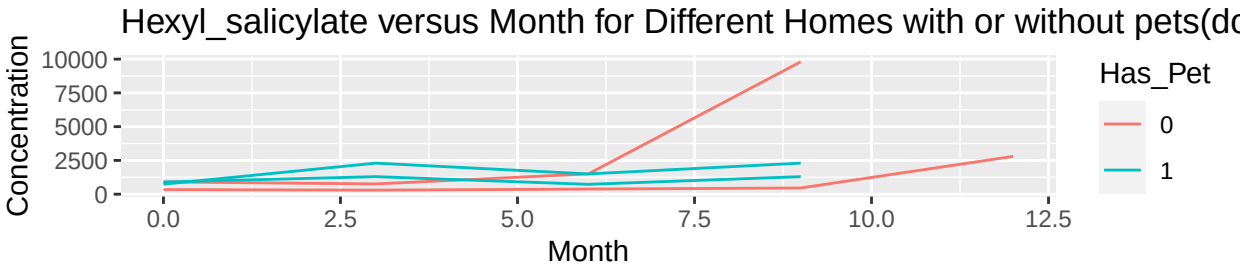

BP versus Month for Different Homes with corresponding usage of paint a
Concentration
30000
20000
10000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Usage of Paints
0
1
Usage of Low VOC Material
−7
0
1

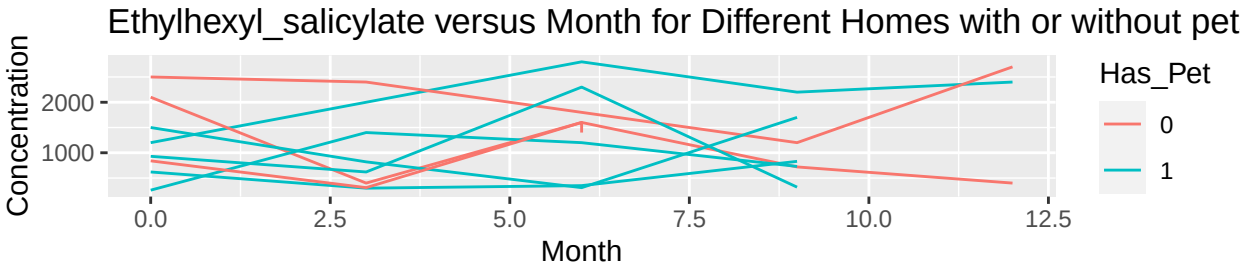

Ethylhexyl_salicylate versus Month for Different Homes with correspondinç
Concentration
2000
1000
0.0
2.5
5.0
7.5
10.0
12.5
Month
Usage of Paints
0
Usage of Low VOC Material
−7
0
1

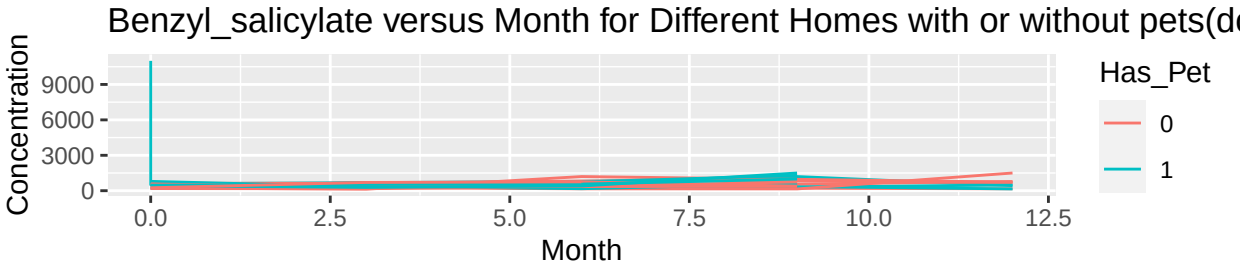

Benzyl_salicylate versus Month for Different Homes with corresponding ty
Concentration
9000
6000
3000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Usage of Low VOC Material
−7
0
1


Benzyl_salicylate versus Month for Different Homes with corresponding us
Concentration
9000
6000
3000
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Usage of Paints
0
1

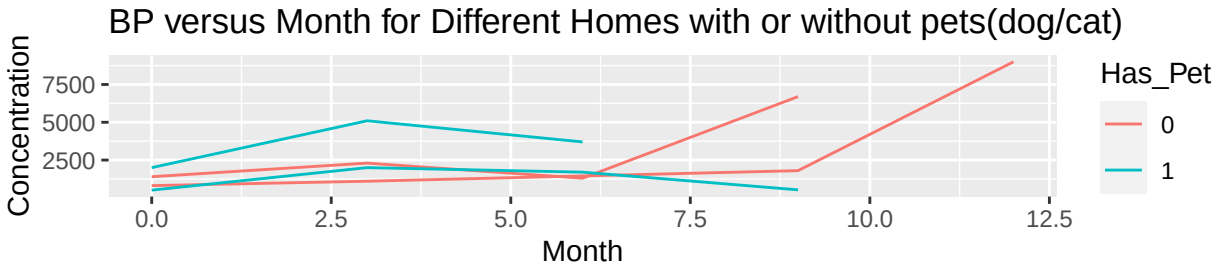

Hexyl_salicylate versus Month for Different Homes with corresponding us
Concentration
10000
7500
5000
2500
0
0.0
2.5
5.0
7.5
10.0
12.5
Month
Usage of Paints
0
Usage of Low VOC Material
−7
0

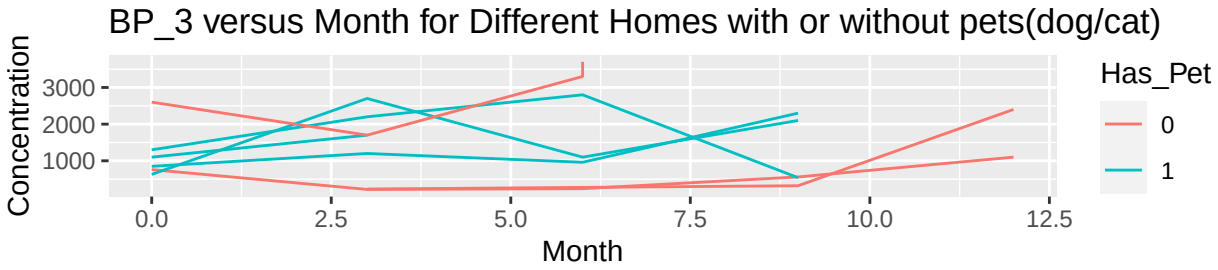


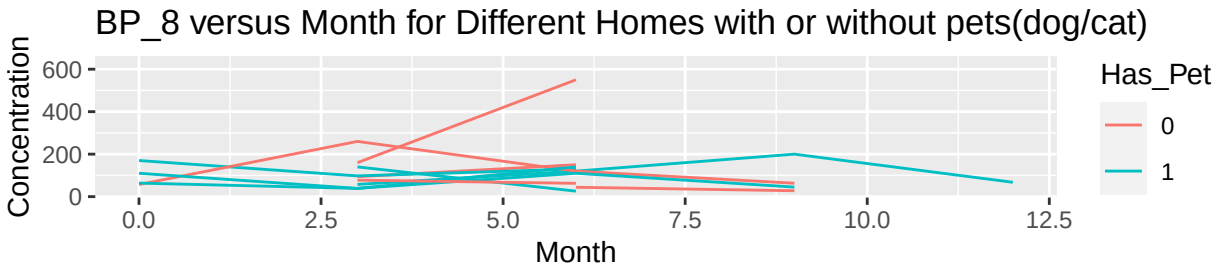





```
## Warning: Removed 1 row containing missing values (`geom_line()`).
## Warning: Removed 4 rows containing missing values (`geom_line()`).
## Warning: Removed 27 rows containing missing values (`geom_line()`).
```

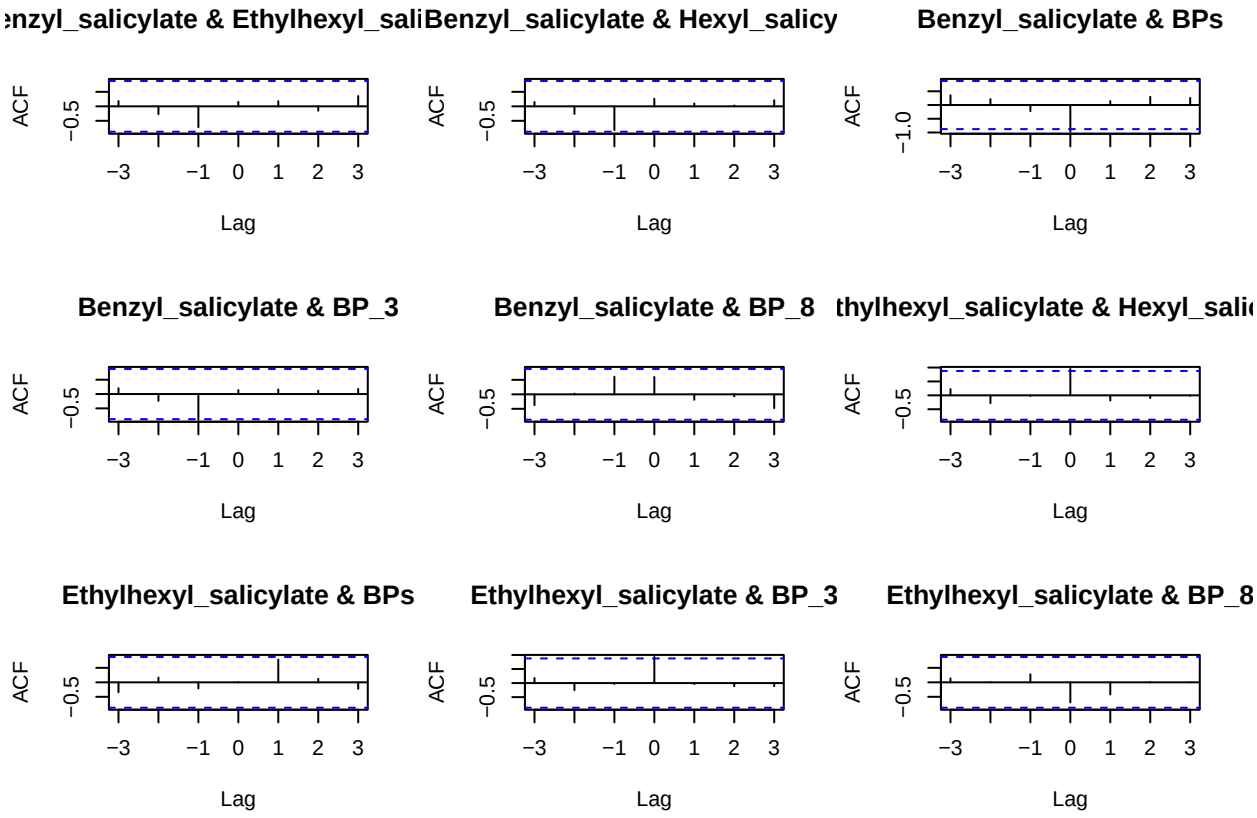




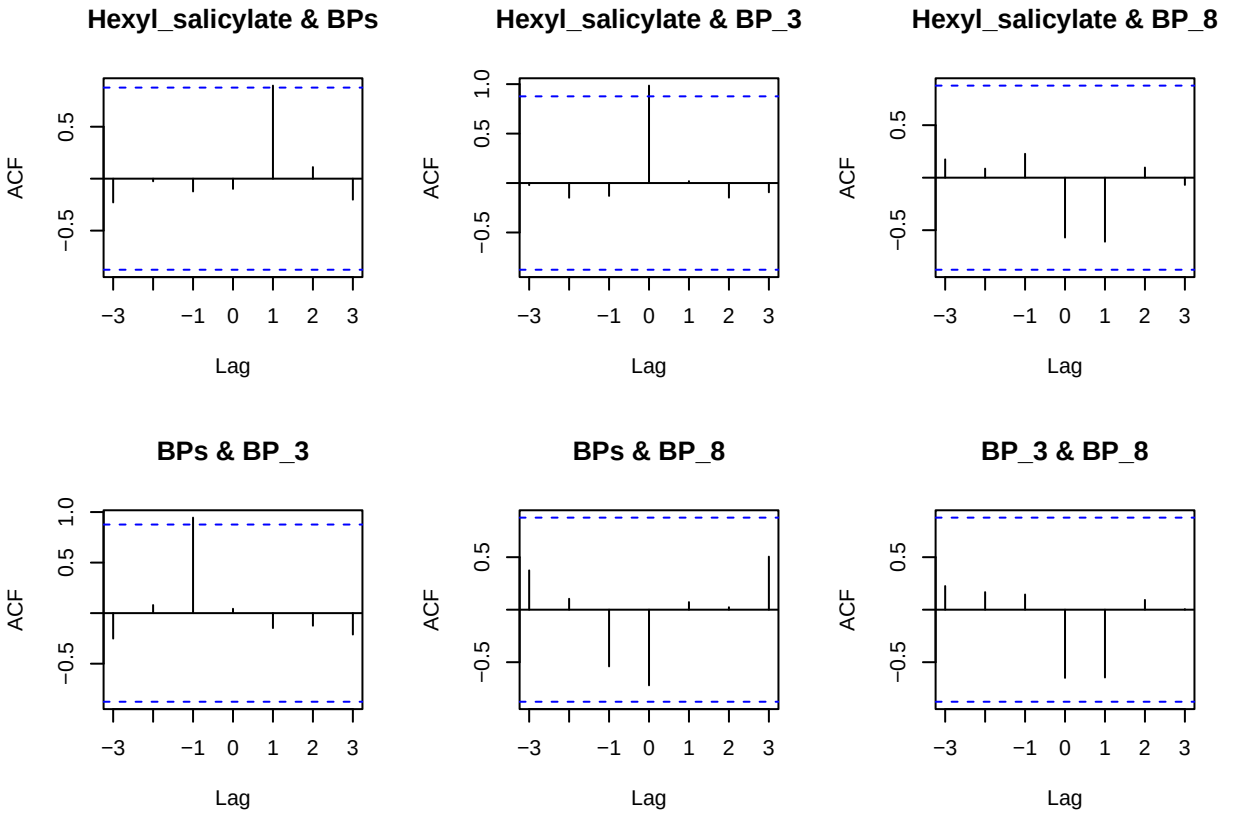



```
par(mfrow = c(3,3))

ccf1(house32)
```




```
par(mfrow = c(2,3))
ccf2(house32)
```



House 29: ES BP3 lag1 BP BP8 lag1

House 30: Hexyl_salicylate and BP_3 has significance in house30 at lag 0

House 31: Hexyl_salicylate and BP_3 has significance in house31 at lag 1

House 32: BS & BPs at lag 0 BS & BP3 at lag -1 ES & HS at lag 0 ES & BP3 at lag -1 HS & BPs lag 1 BPs & BP3 lag -1 HS & BP3 lag 0

This suggests that for house32, many chemicals have similar trend almost at the same time or at a slightly different time.

House 34: ES BP8 lag0 BP3 BP8 lag0 HS BP3 lag0 HS BP8 lag0

House36: BS HS lag0

Positive correlations at lag 0 or 1 suggest a direct or short-term positive relationship between the variables, indicating that they tend to increase or decrease together, either simultaneously or with a short delay.

Negative correlations at lag 0 or 1 suggest an inverse relationship, indicating that an increase in one variable is associated with a decrease in the other, either simultaneously or with a short delay.