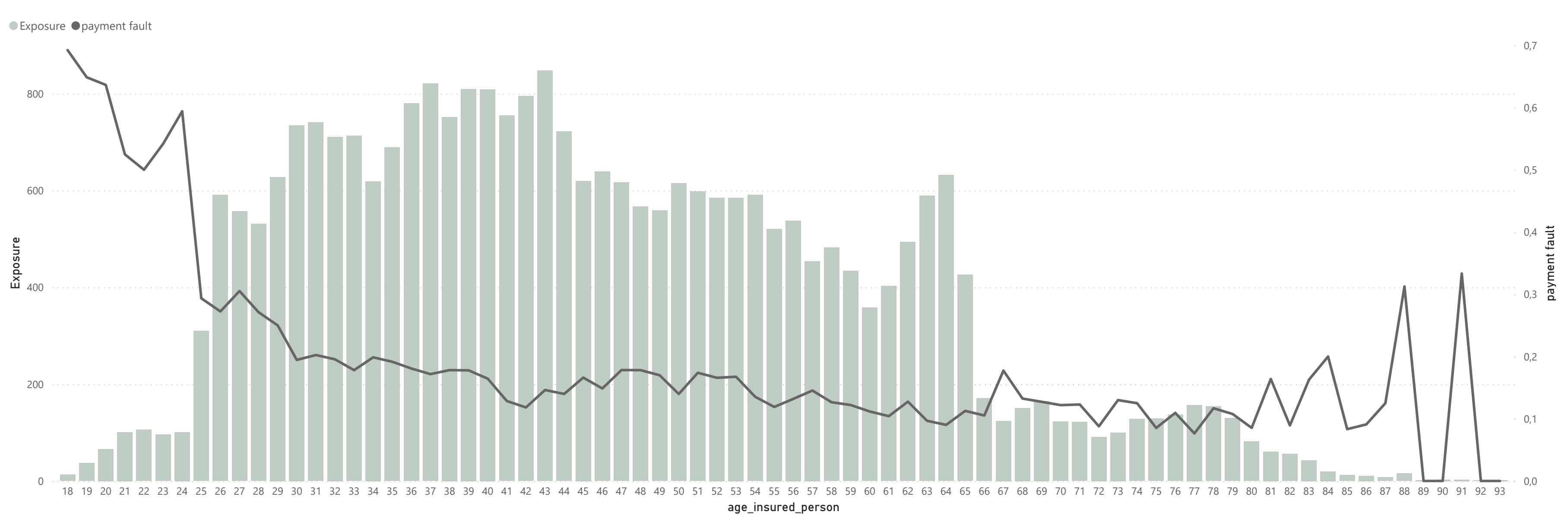
Quick view the Data

(I) Quickview on Data

- Get a first idea, which fields might have a an impact on payment faults
- Find areas with weak exposure and prepare binning and grouping values
- Check for interactions between fields

Preview the data regarding a payment fault model

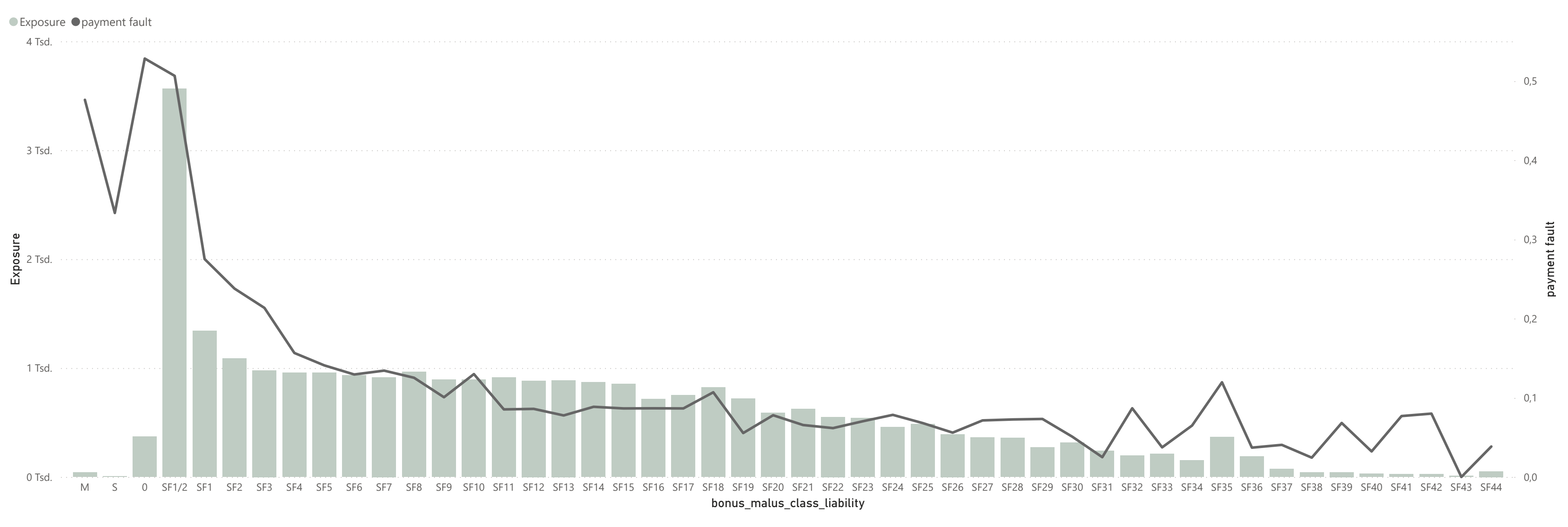
Age of insured person



(i) there is definitely some correlation between young drivers and payment faults. The significance might not hold, due to the small exposure

- Age groups are probably good. especially under 25, over 80 and some sections in the middle
- Maybe it is enough to go young Y/N

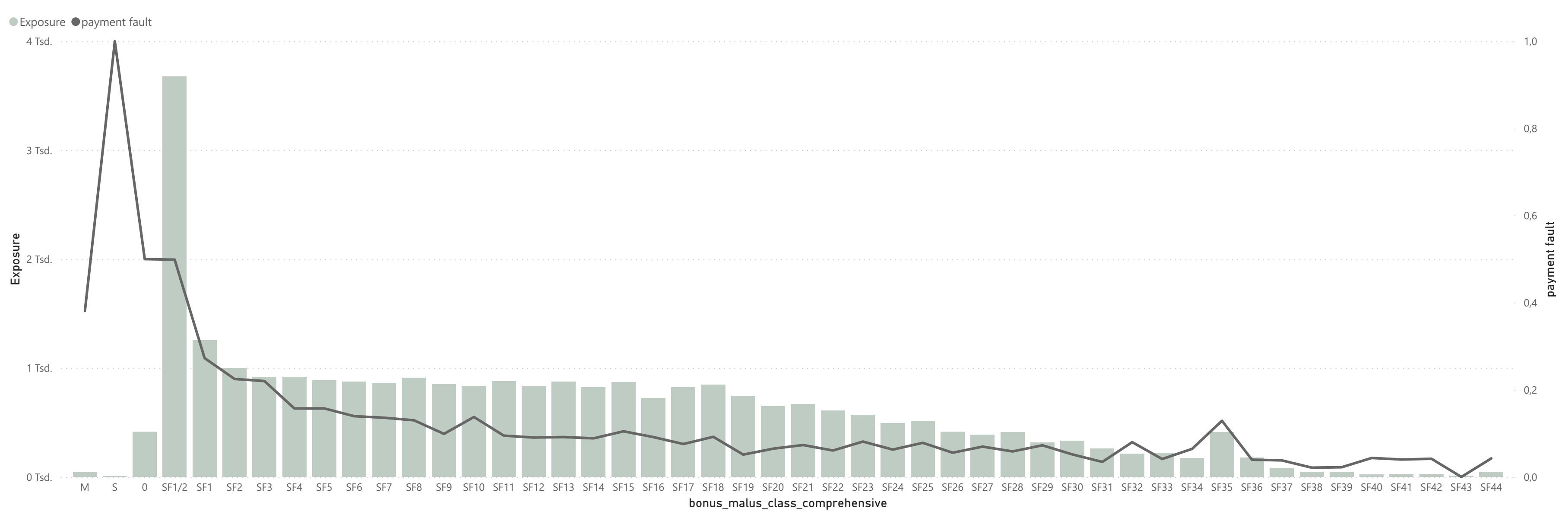
SF-Class third party liability



(i) higher price correlates strongly with the payment_fault

- Group [: 1/2][1:4][5:10][11:]
- Check correlation with young drivers

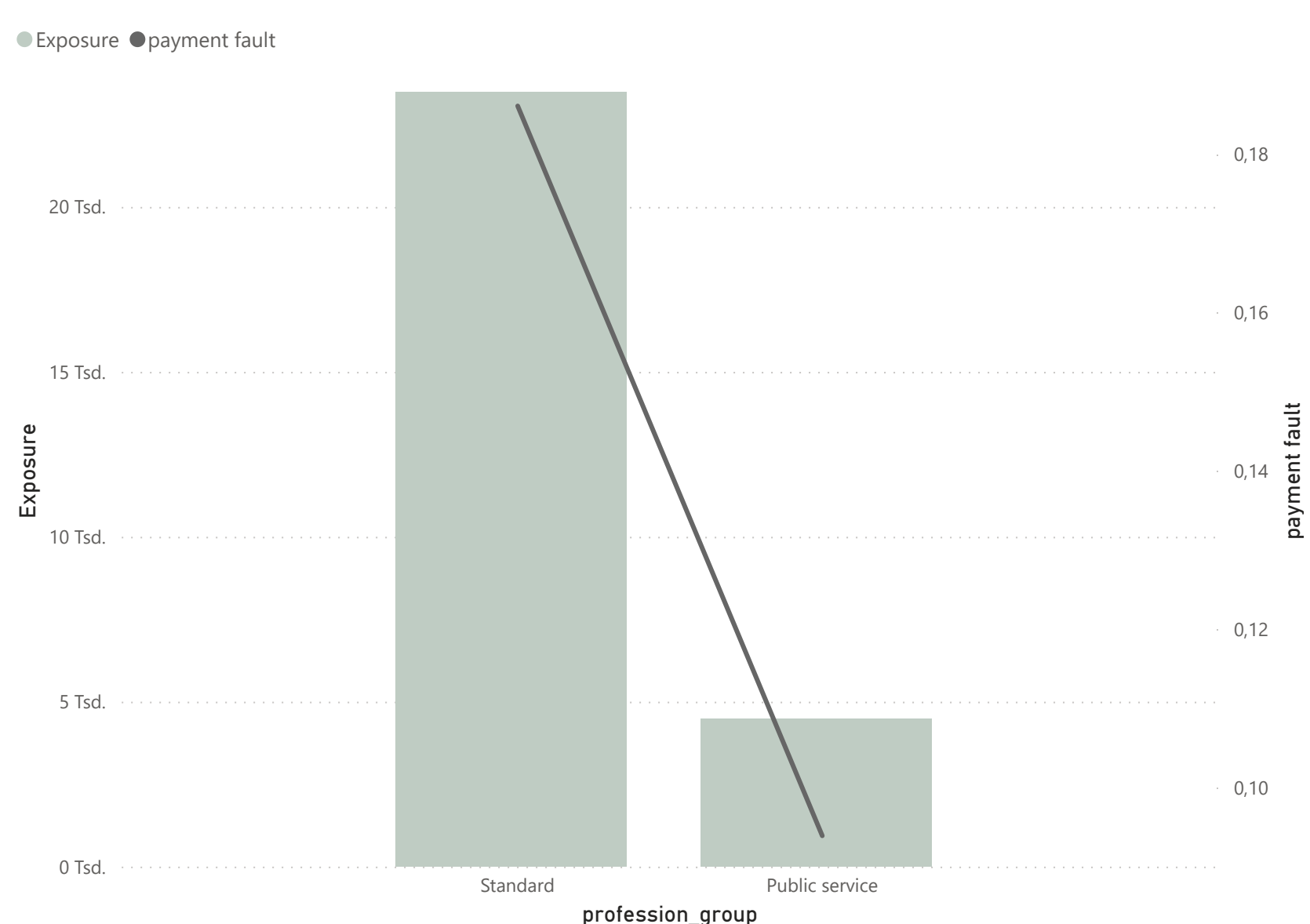
SF-Class full comprehensive cover



(i) don't put this in, becaus:

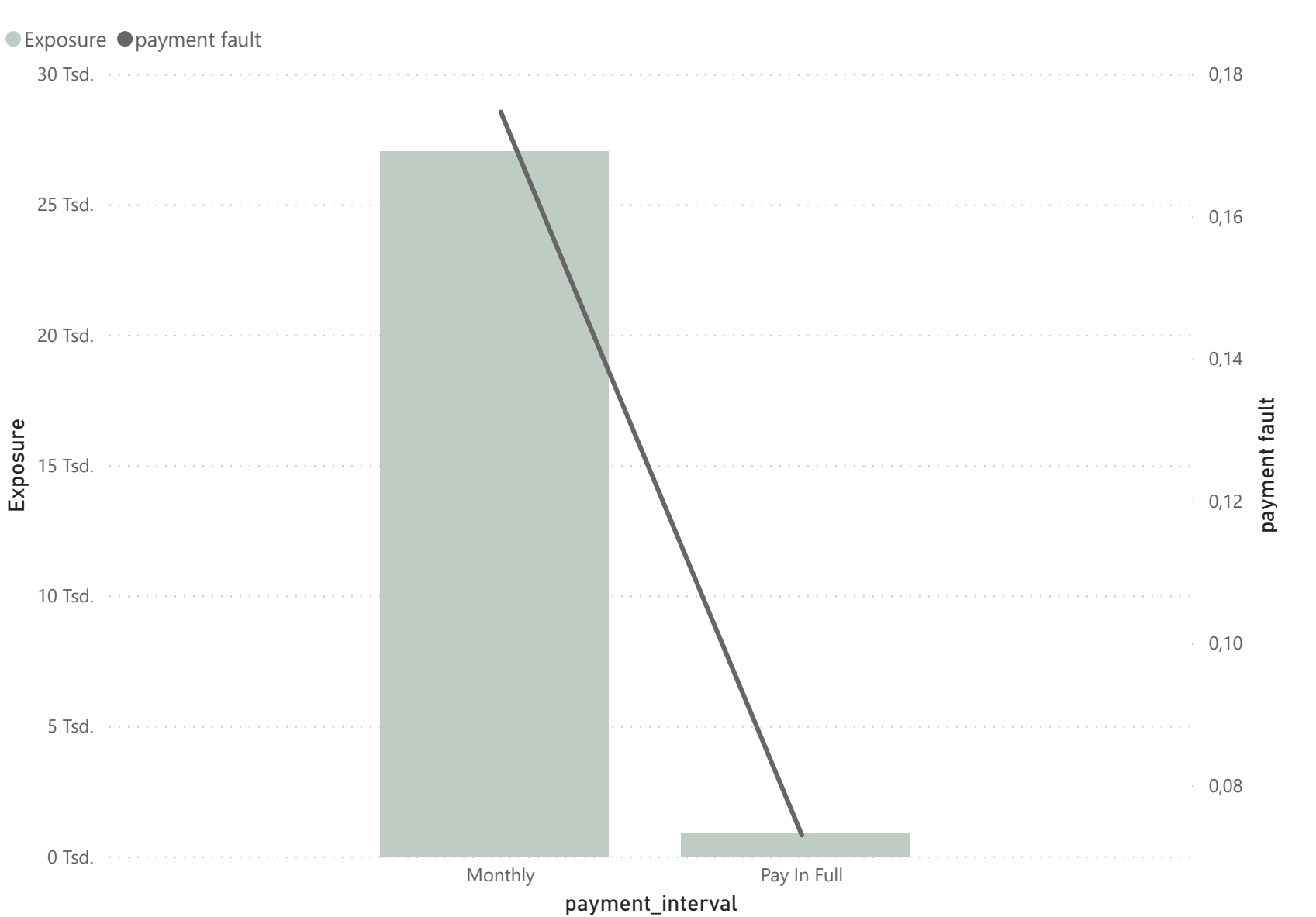
- high correlation with SF_tpl
- only a fraction has it and Cover VK is very low compared to the others

Profession group



(i) The profession group is either a diffrent type of customer, or the pricing is very different.

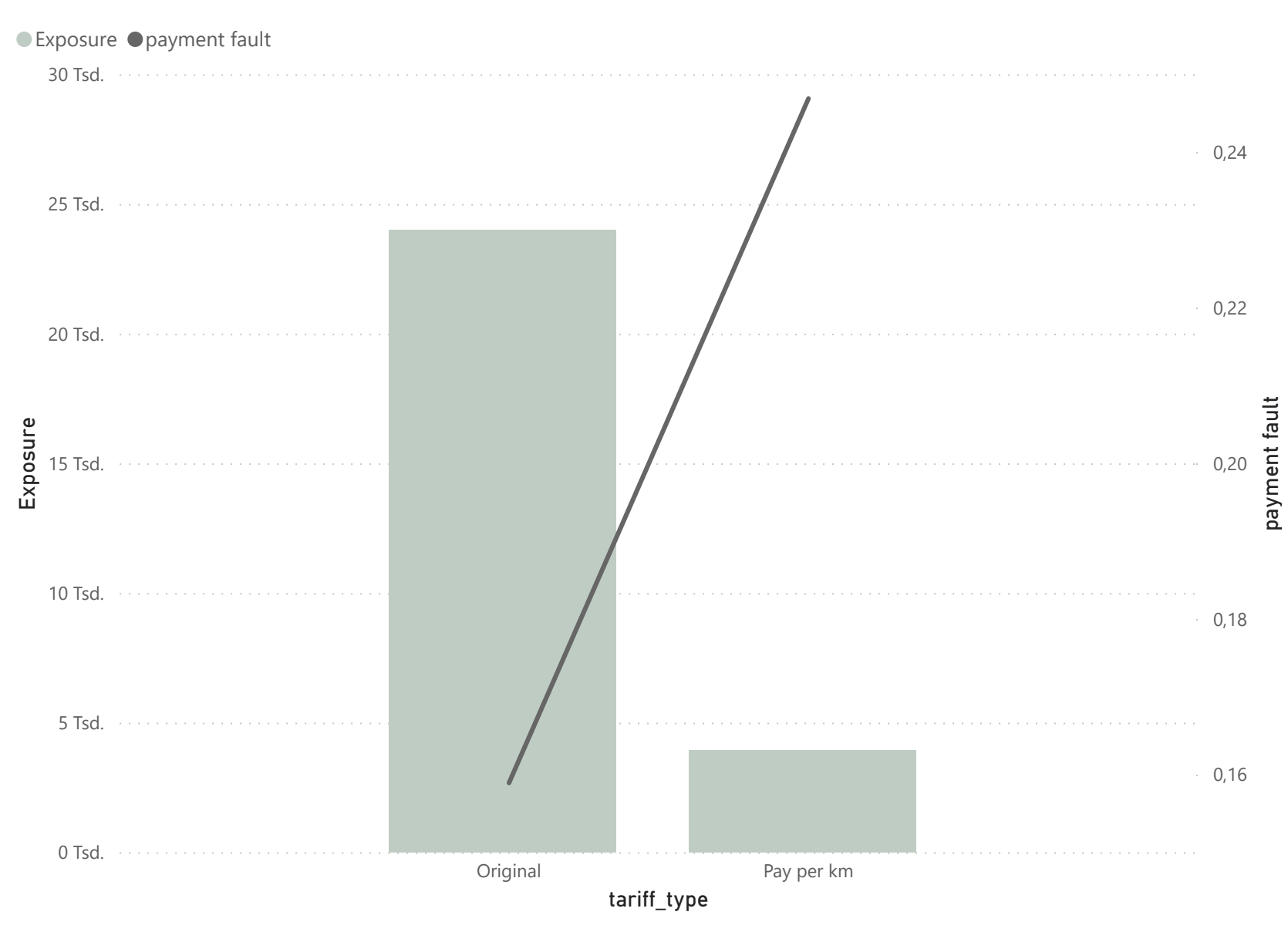
Payment



(i) This seems very strong

- use this as is but also with interactions

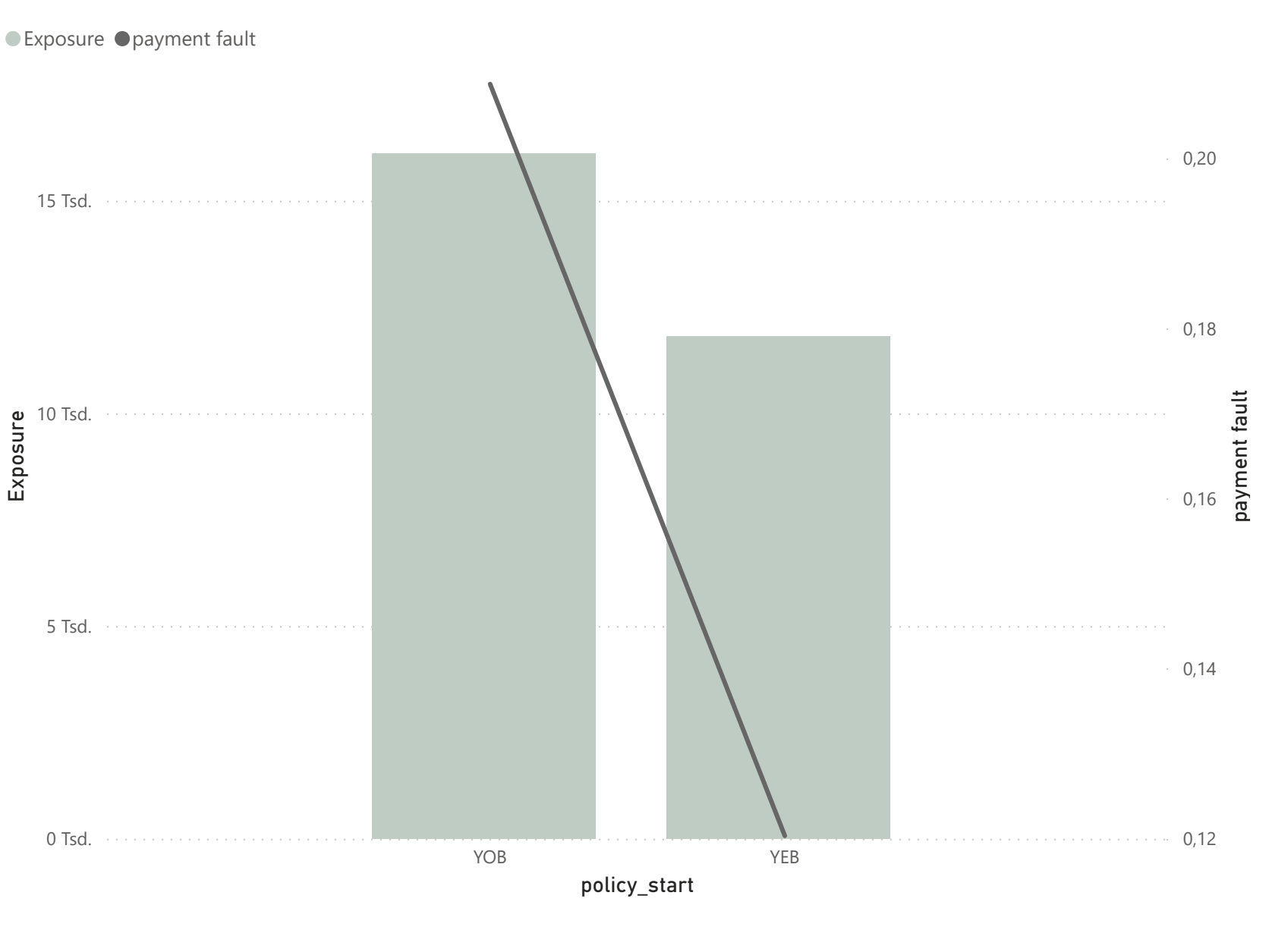
Tarif type



(i) The Tarif has a influence on the payment fault

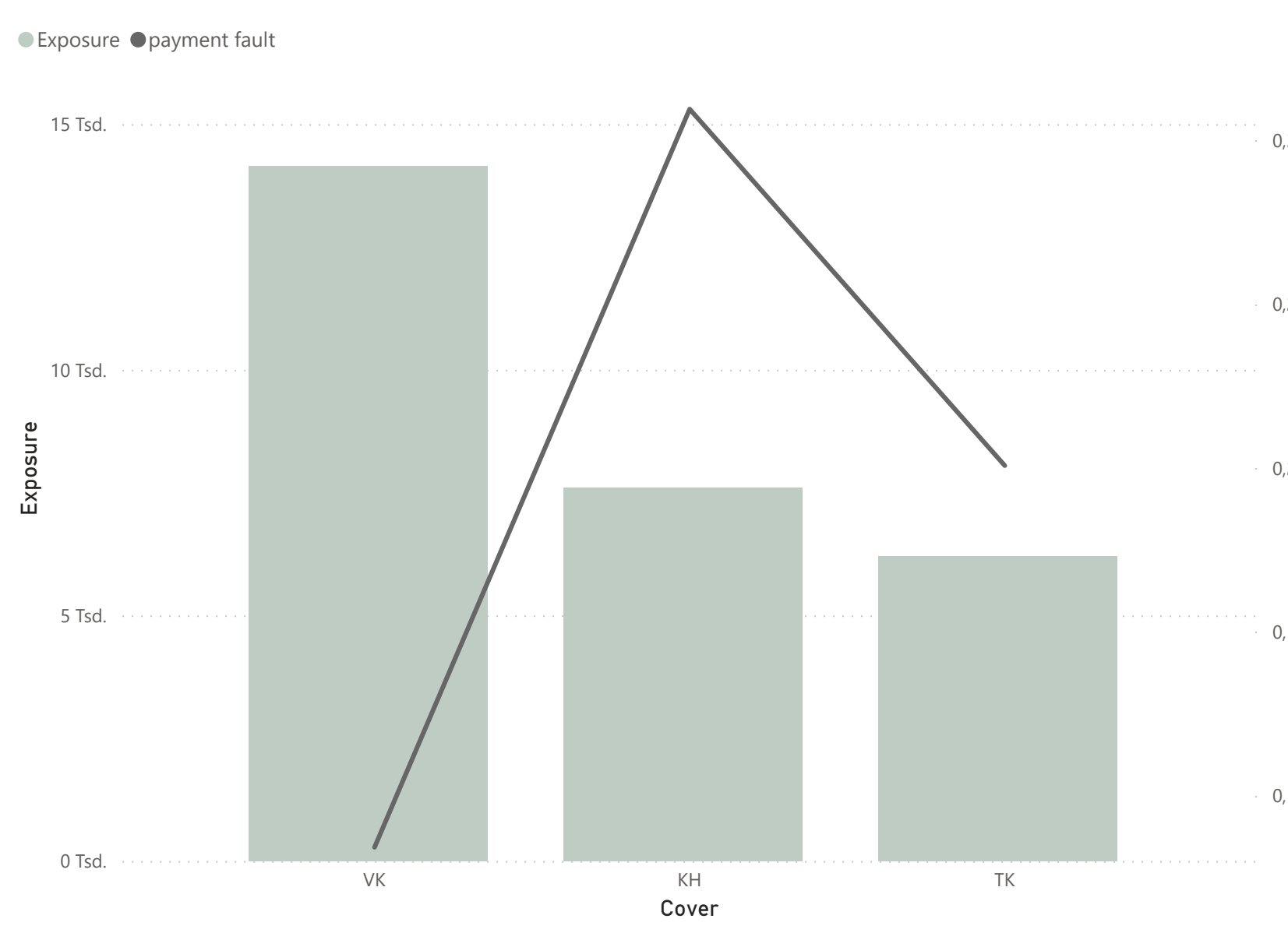
- There seem to be little correlation to the other fields (check that again in python)
- as seen above: if the contracts are "payed in Full" there is no little to no between the classes

Policy start



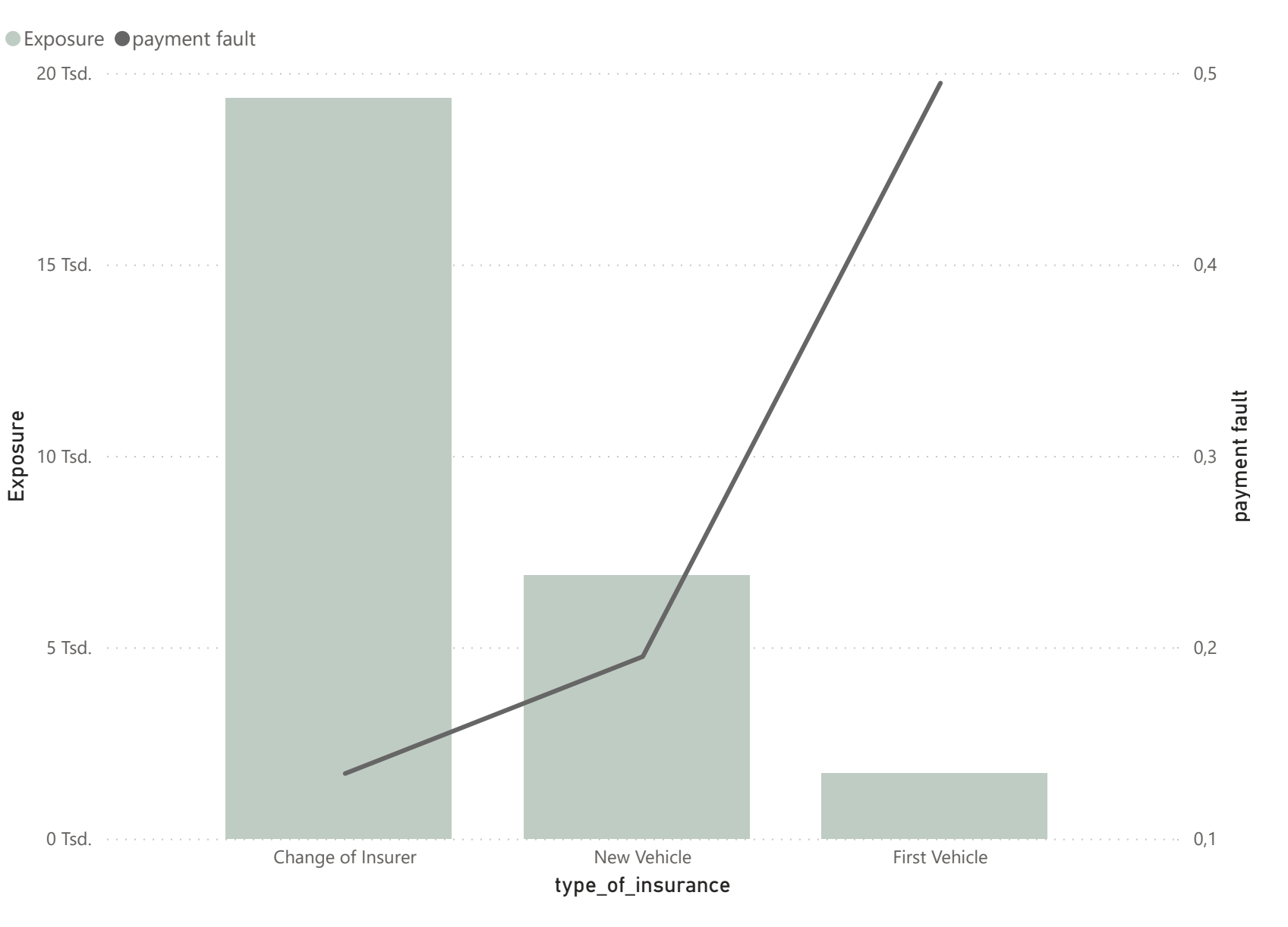
(i) correlates strongly with Friday first insurer

Cover



(i) definitly relevant

Type of insurence

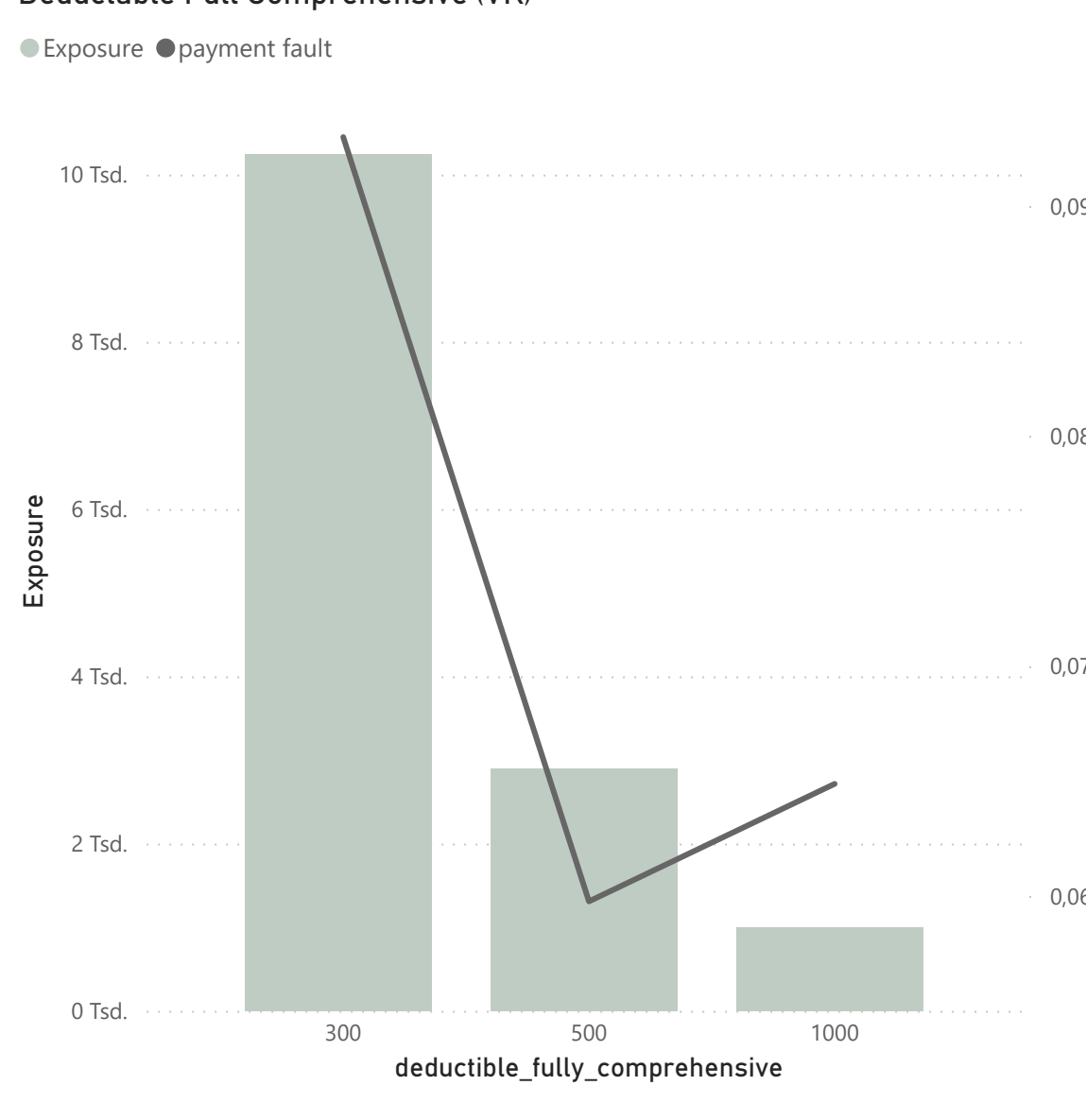


(i) The tariff type has a influence on the payment fault

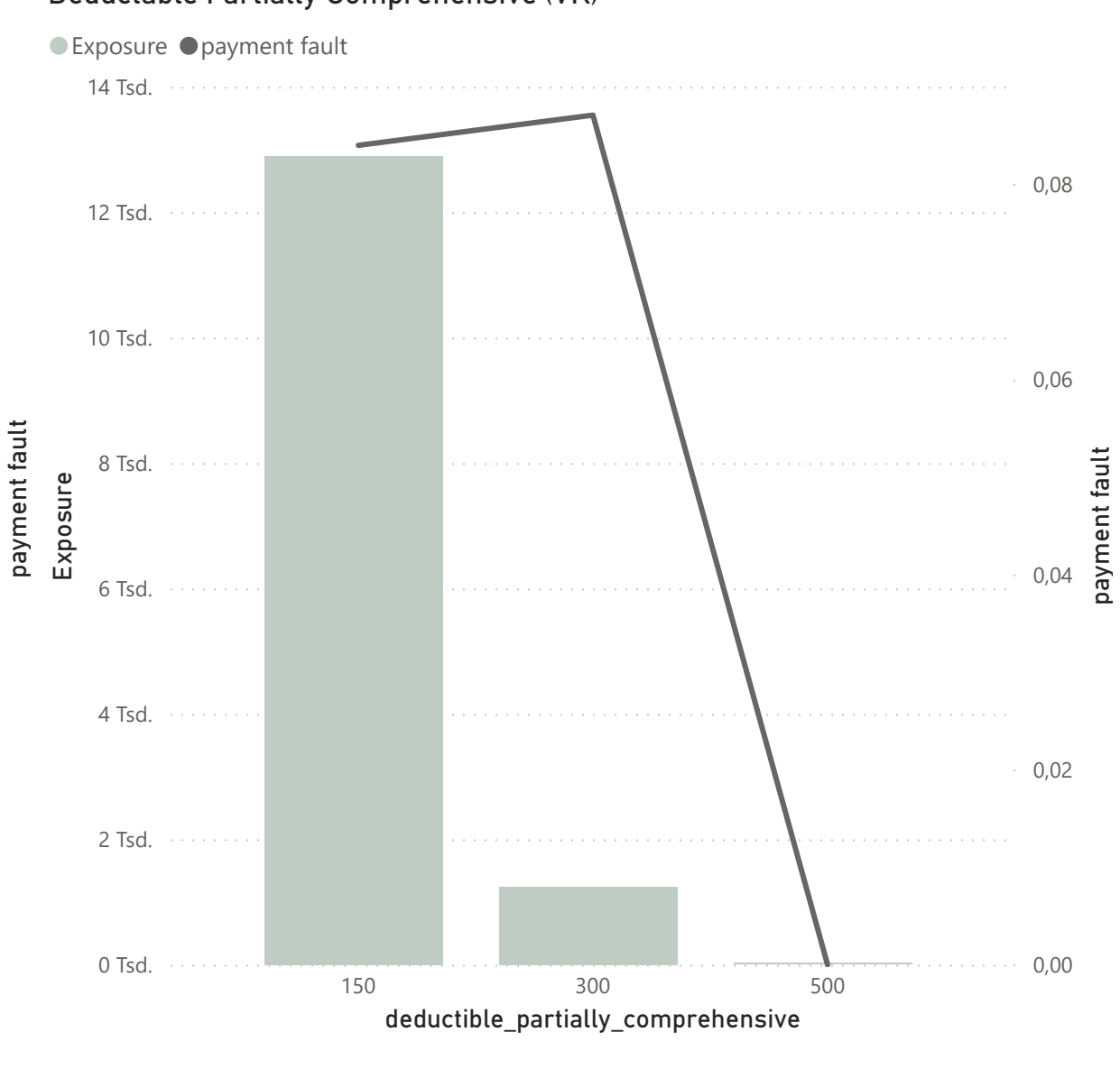
- There seem to be little correlation to the other fields (check that again in python)
- as seen above: if the Contracts are "payed in Full" there is no little to no between the classes

Deductibles

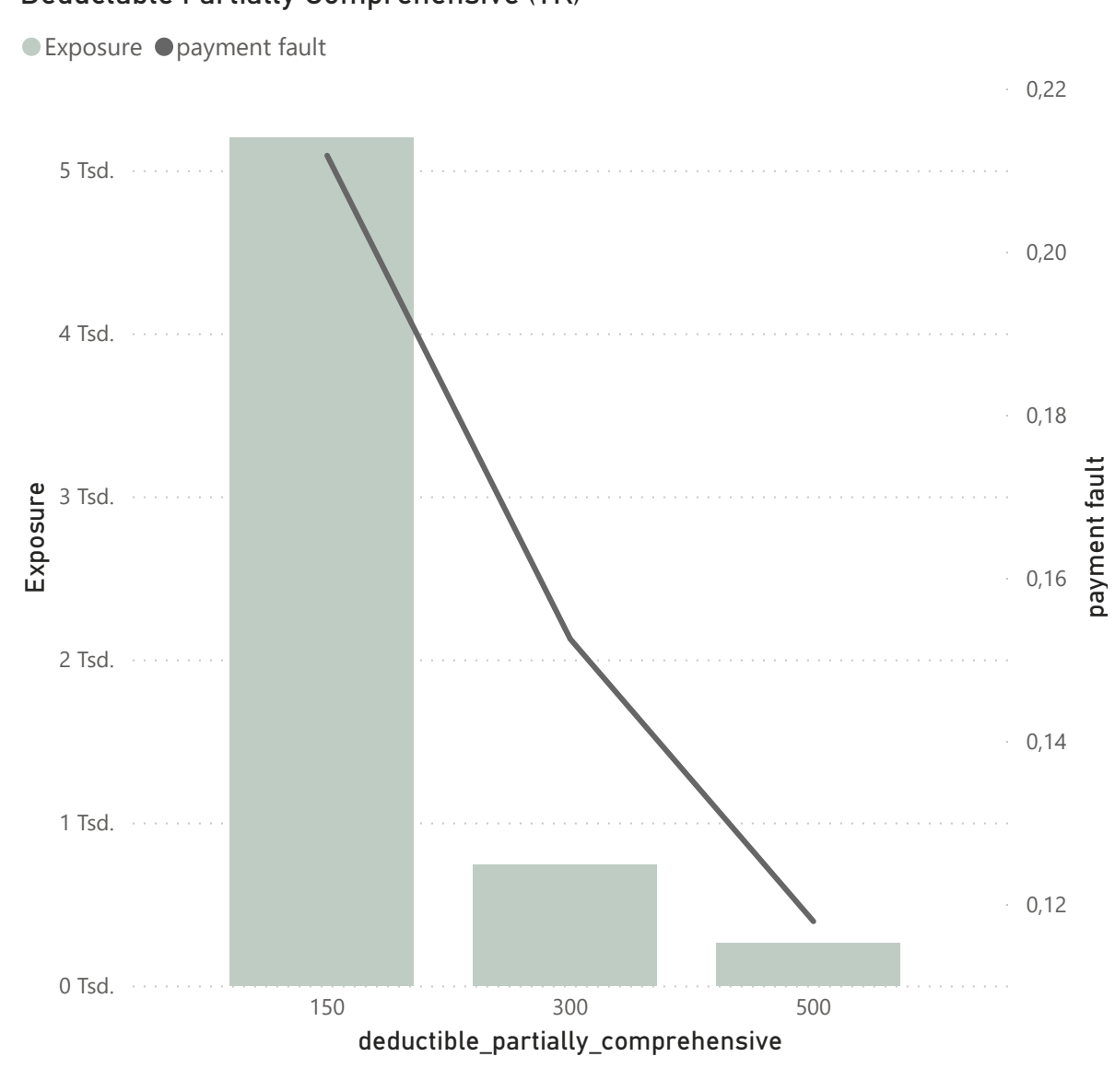
Deductable Full Comprehensive (VK)



Deductable Partially Comprehensive (VK)



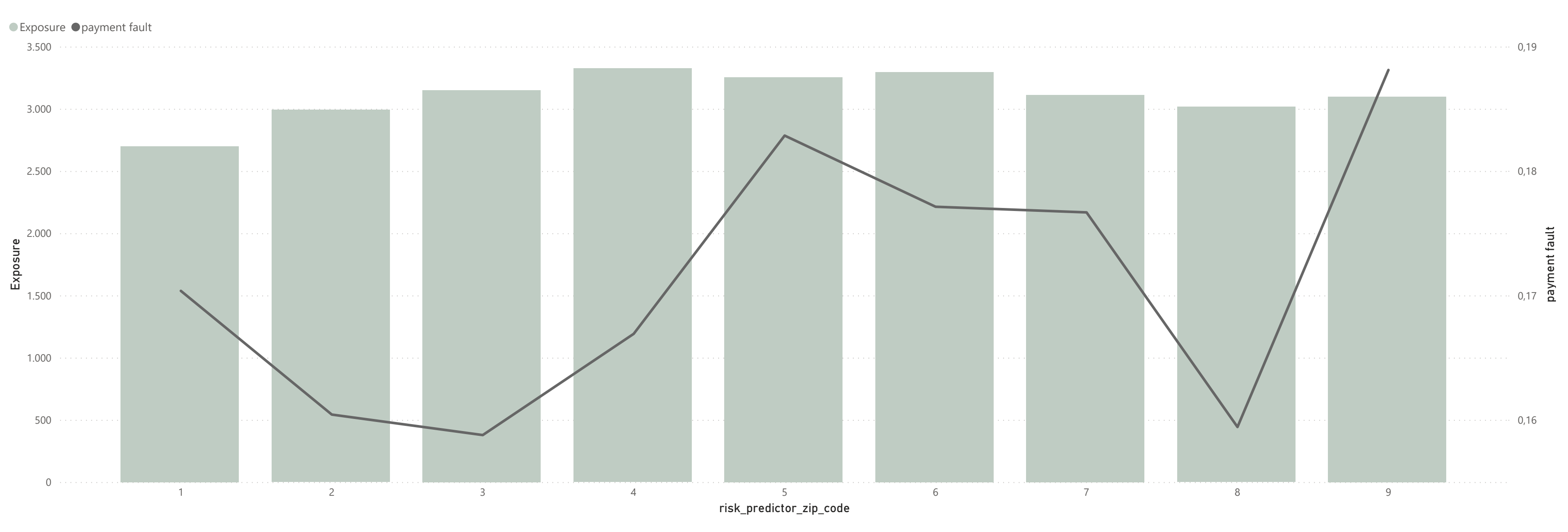
Deductable Partially Comprehensive (TK)



(i) there is defenetly some correlation between young drivers and payment faults. The significants might not hold, due to the small exposure

- in FC. Maybe group 500 and 1000 to get it smooth

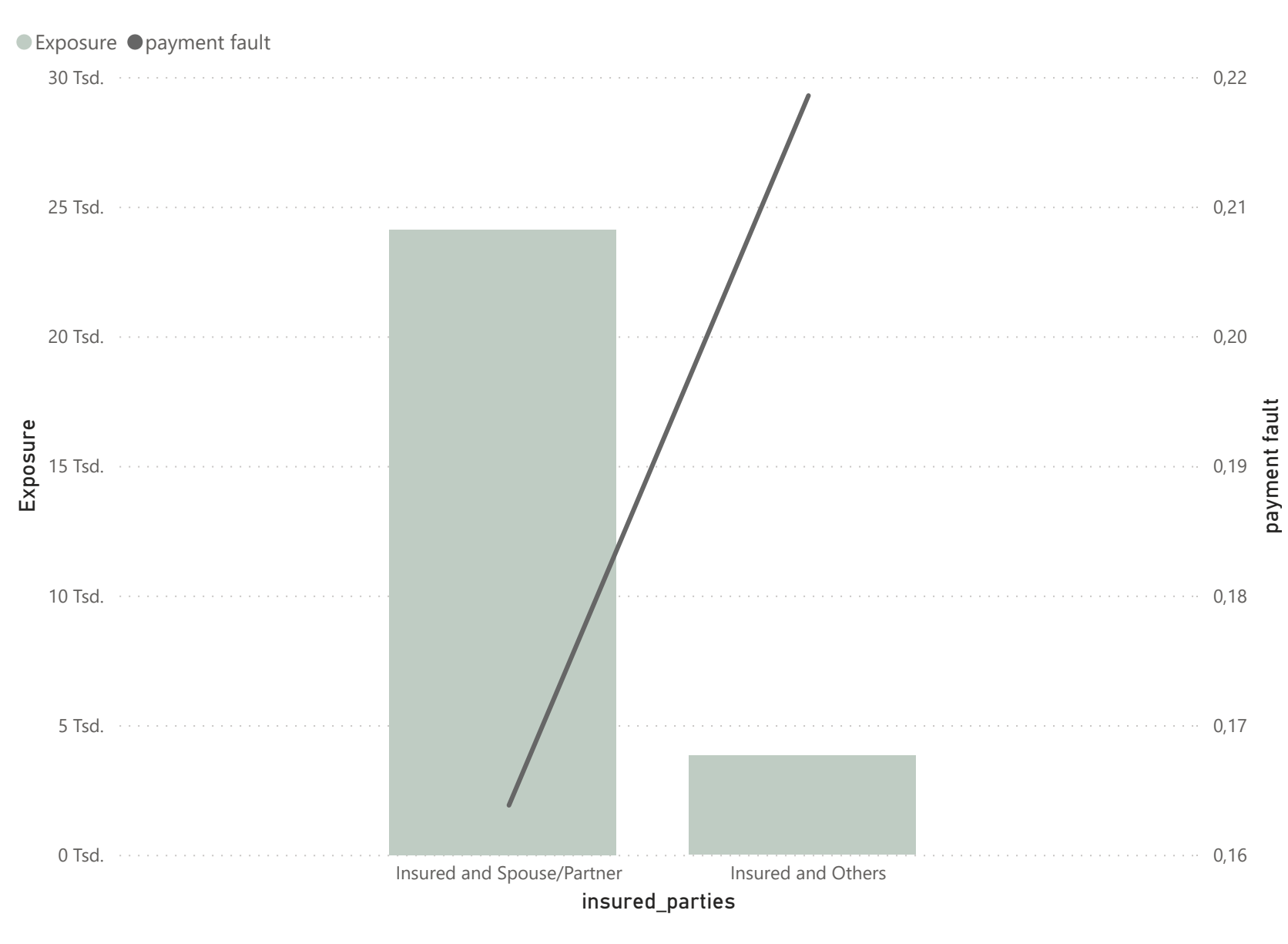
Regional class (FC)



(i) Group MULTIHOTONE:

- [1,4],[2,3,8], [5,6,7],[9]
- **not the strongest field**
- **There is no different risk_predictor_zip_codes for all the different coverages**

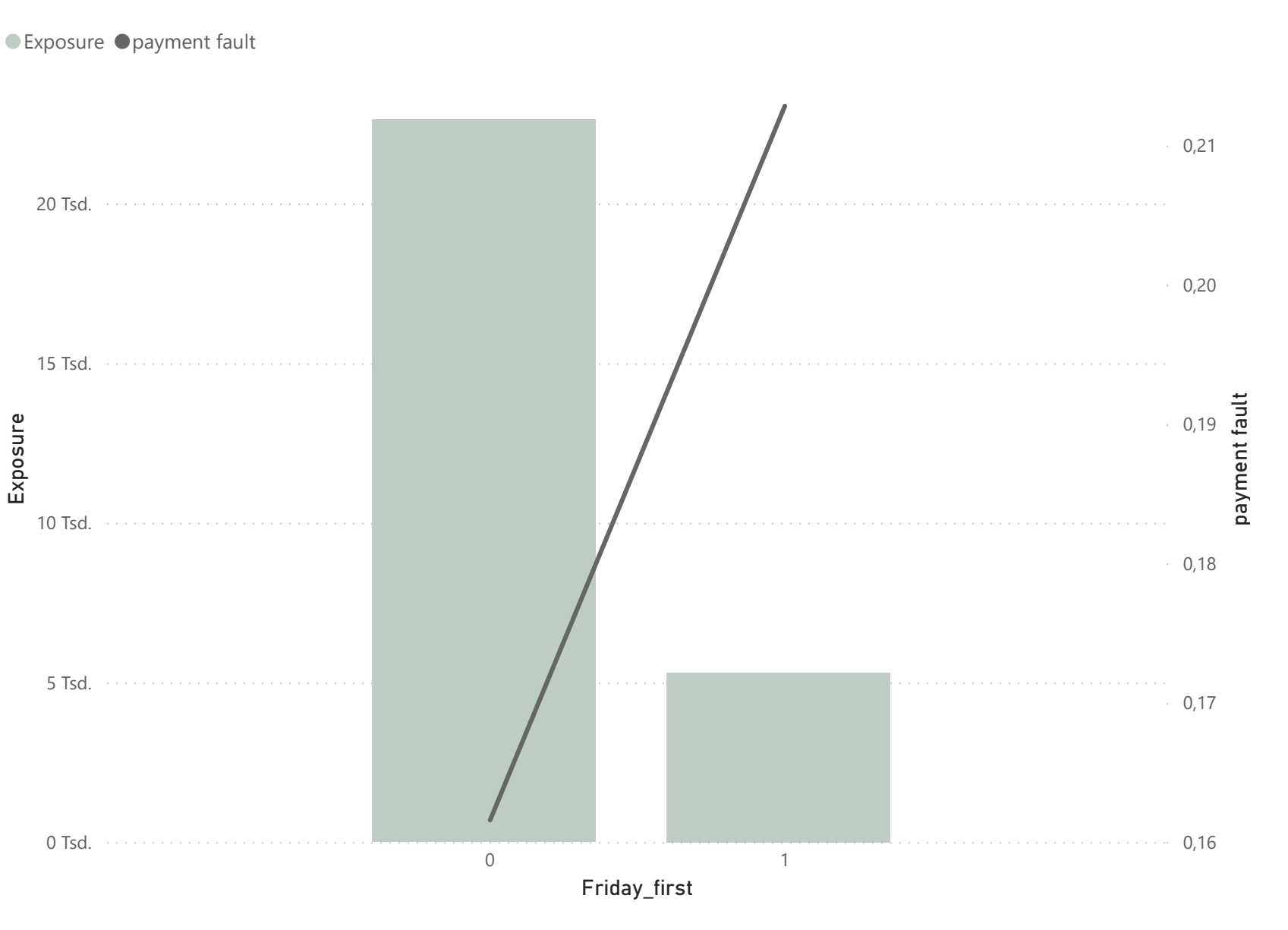
Insured parties



(i) there is a significant differance

- check for interactions later

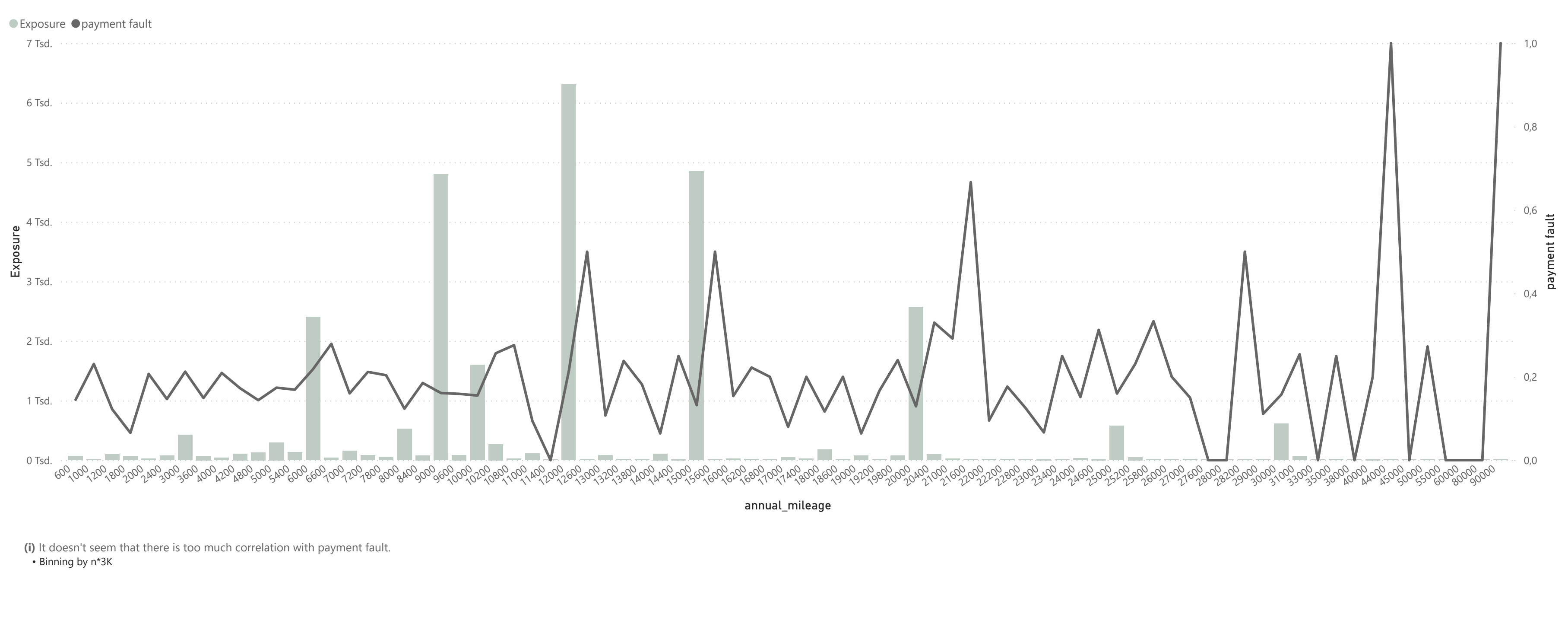
Friday first insurer (car age purchase = contract start)



(i) Friday first insurer
(i) sadly the rate is much higher in the group where Friday is the first insurer after purchasing

- this correlates highly with the bm_class 1/2
- this correlates highly Policy Start

Annual mileage

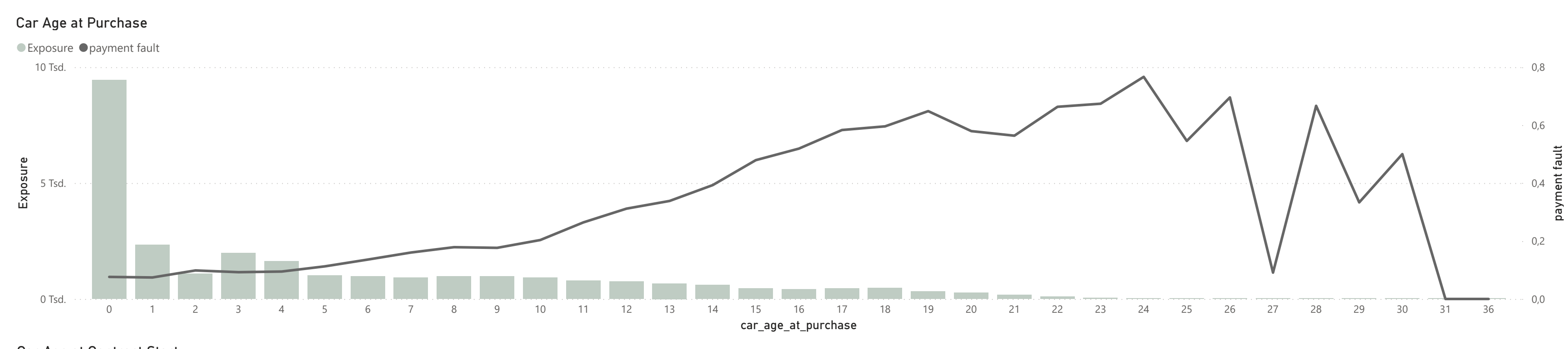


(i) It doesn't seem that there is too much correlation with payment fault.

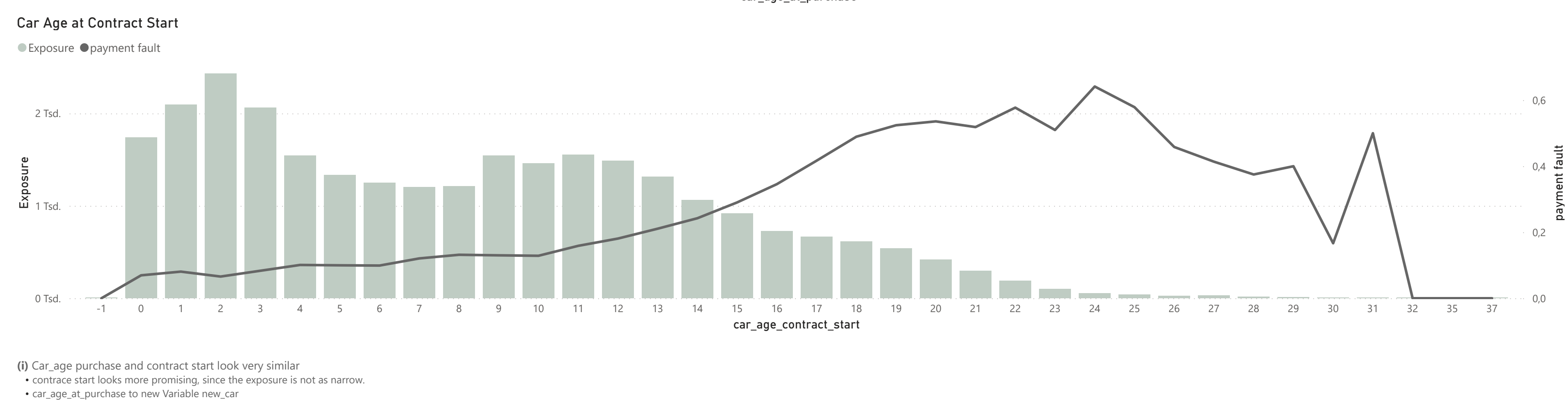
- Binning by n*3K

Car age

Car Age at Purchase



Car Age at Contract Start

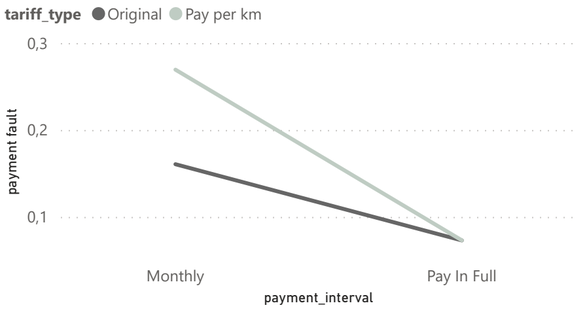


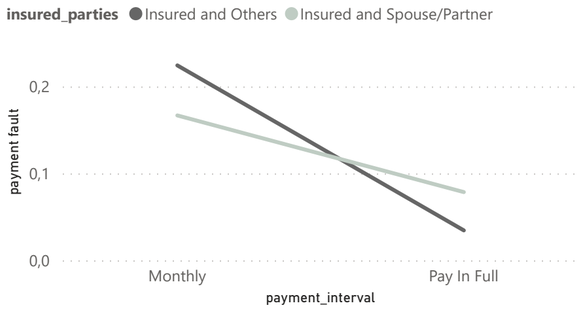
(i) Car_age purchase and contract start look very similar

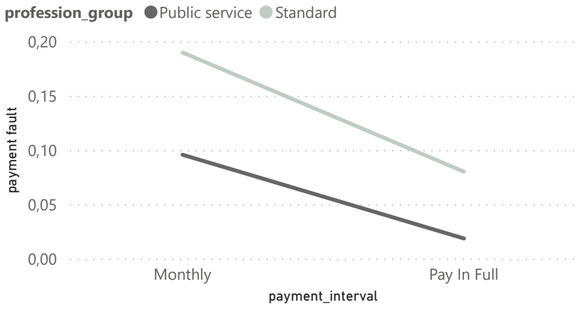
- contrace start looks more promising, since the exposure is not as narrow.
- car_age_at_purchase to new Variable new_car

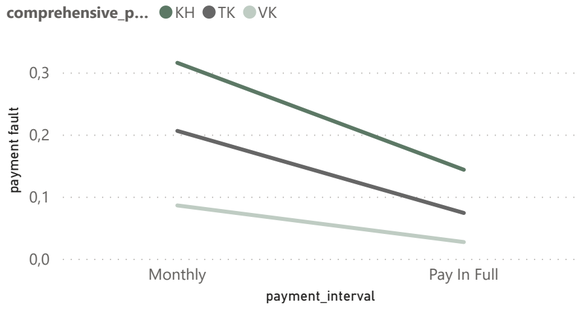
Test Data for Interactions

Possible interactions with Payment type

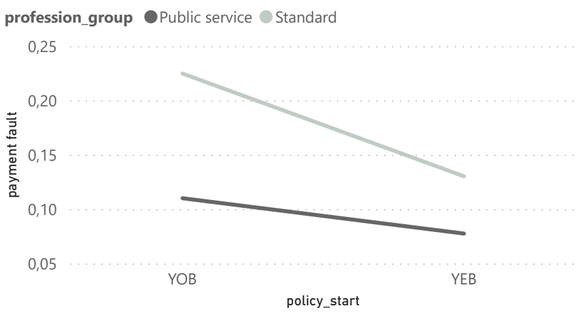




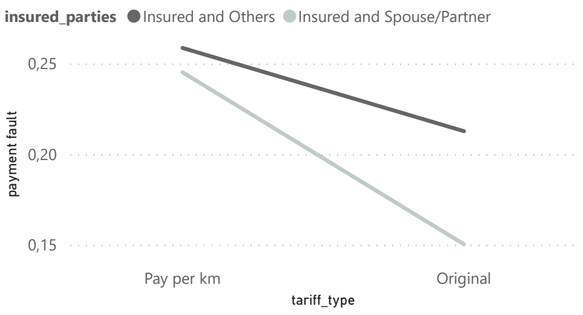




Possible interactions with policy start



Possible interactions with tariff type



Summary: quick view the Data

(I) Pricing:
As expected, payment faults seem to correlate with a higher price tag. This shows by a high correlation between fields that usually have a greater impact in pricing models.

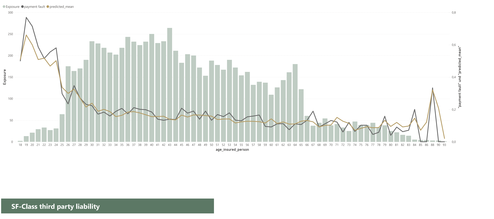
(II) Customer type
But there is another component, the type of customer. It seems, that the profession, the number of installments as well as the (usually relative expensive) full comprehensive coverage does have a positive impact on payment faults.

This makes sence, because price sensitive customers tend not to buy such a coverage (porbably because the car is not expensive enough as well).

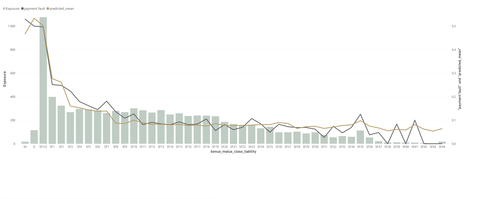
Notes:
the given regional classes are for full comprehensive cover and typically do not affect the pricing of a third liability or partially comprehensive covers. So therefore the effect would be stronger if the regional classes for third liability would have been given (even though they correlate).

Evaluating fit on test data only

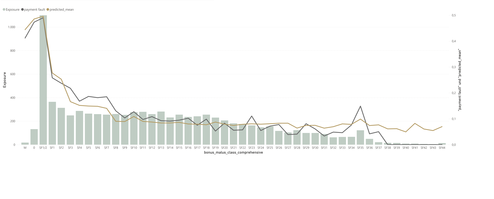
Age of insured person



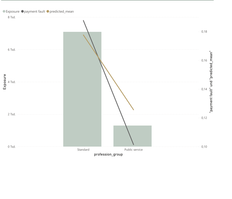
SF-Class third party liability



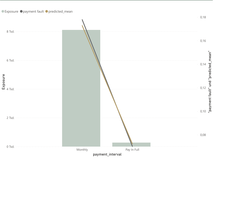
SF-Class full comprehensive cover



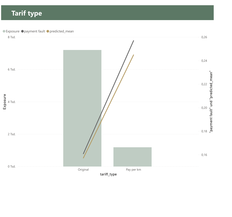
Profession group



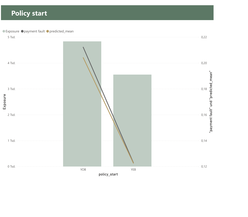
Payment



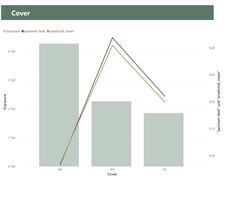
Tarif type



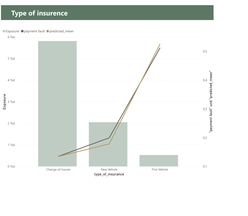
Policy start



Cover

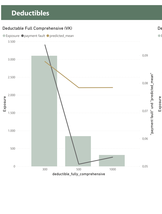


Type of insurence

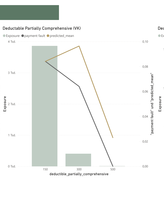


Deductibles

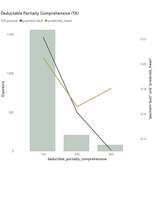
Deductable Full Comprehensive (VK)



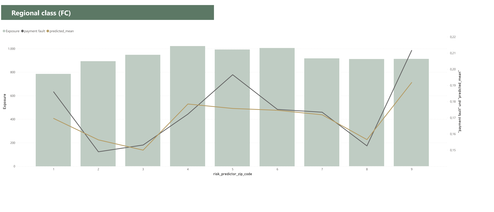
Deductable Partialy Comprehensive (VK)



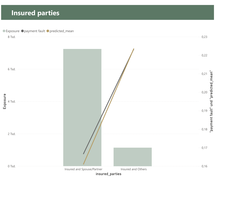
Deductable Partialy Comprehensive (TK)



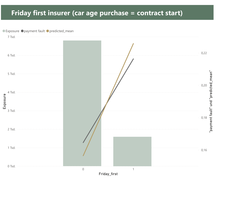
Regional class (FC)



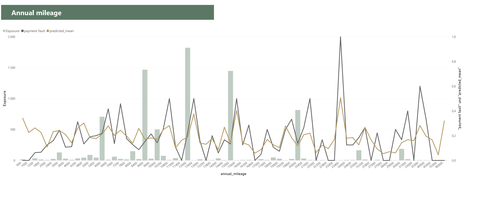
Insured parties



Friday first insurer (car age purchase = contract start)

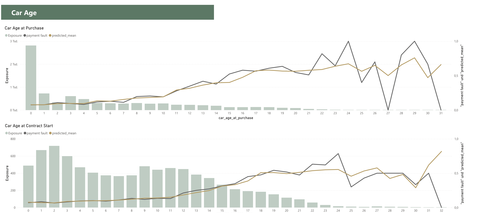


Annual mileage

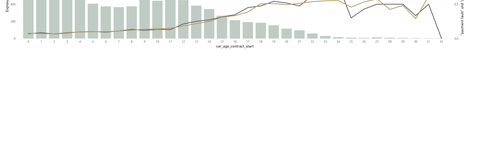


Car Age

Car Age at Purchase



Car Age at Contract Start



Next Steps:

(I) Checking, that all used Features are actually available at the point of the application

(II) Extracting the factors and bounds for each field to generate a rating matrix

(III) Testing the rating matrix for holes