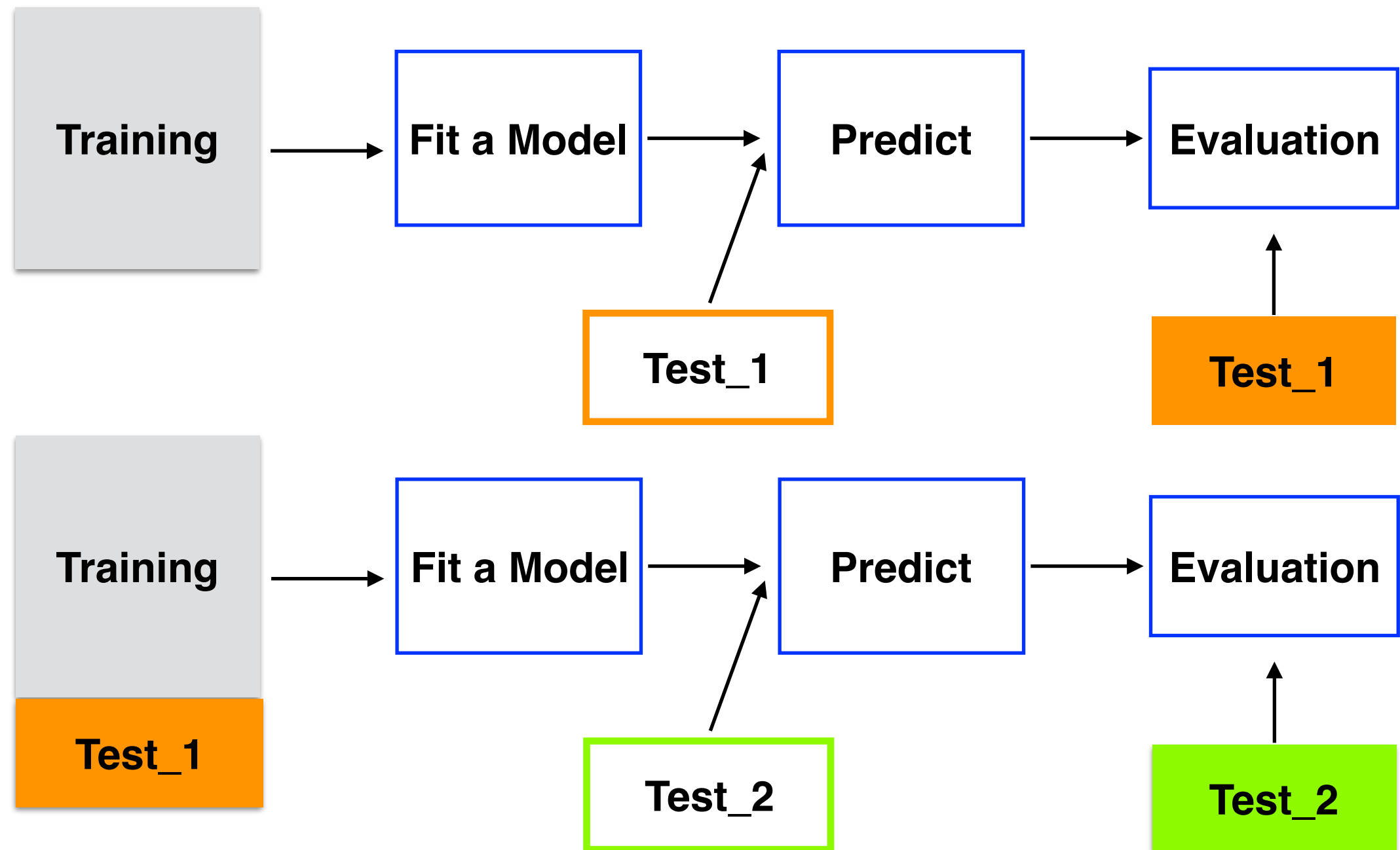


Project 2



Load all data; focus on fold 8, dept = 5

```
old.train = readr::read_csv('train.csv')
t=8
start_date = ymd("2011-03-01") %m+% months(2 * (t - 1))
end_date = ymd("2011-05-01") %m+% months(2 * (t - 1))

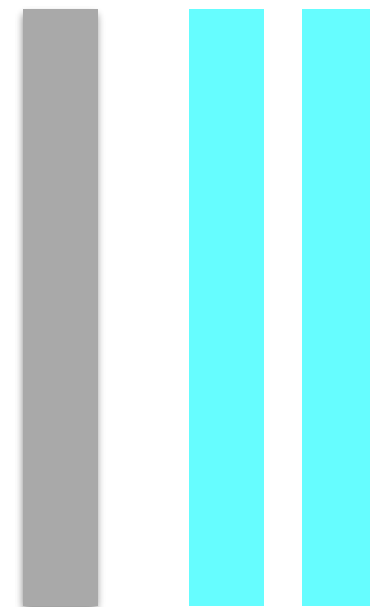
all_depts = unique(old.train$Dept)
dept = all_depts[5]
```

```
train_dept_ts = old.train %>%
  filter(Date < start_date) %>%
  filter(Dept == dept) %>%
  select(Store, Date, Weekly_Sales) %>%
  spread(Store, Weekly_Sales)

dim(train_dept_ts)
#[1] 117 46
range(train_dept_ts$Date)
# [1] "2010-02-05" "2012-04-27"
diff(range(train_dept_ts$Date))
# Time difference of 812 days
```

train_dept_ts

Date S1 S2 S45



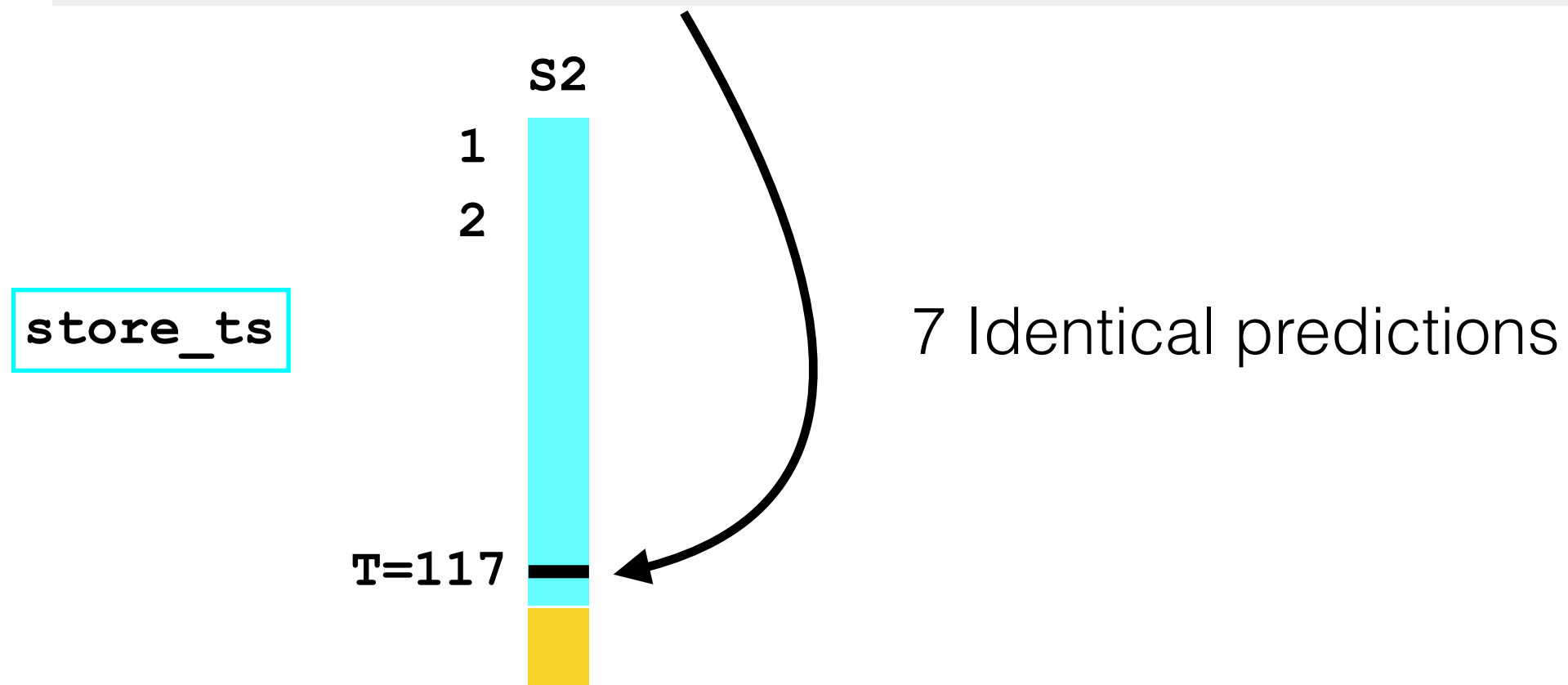
117 rows; 46 cols

Focus on store 2; predict the next 7 wks

```
j=3 #Store 2
store_ts = ts(train_dept_ts[, j], frequency=52)
T = length(store_ts)
num_forecasts = 7
```

1. Naive

```
naive.pred = naive(store_ts, num_forecasts)$mean
store_ts[T]
naive.pred
```

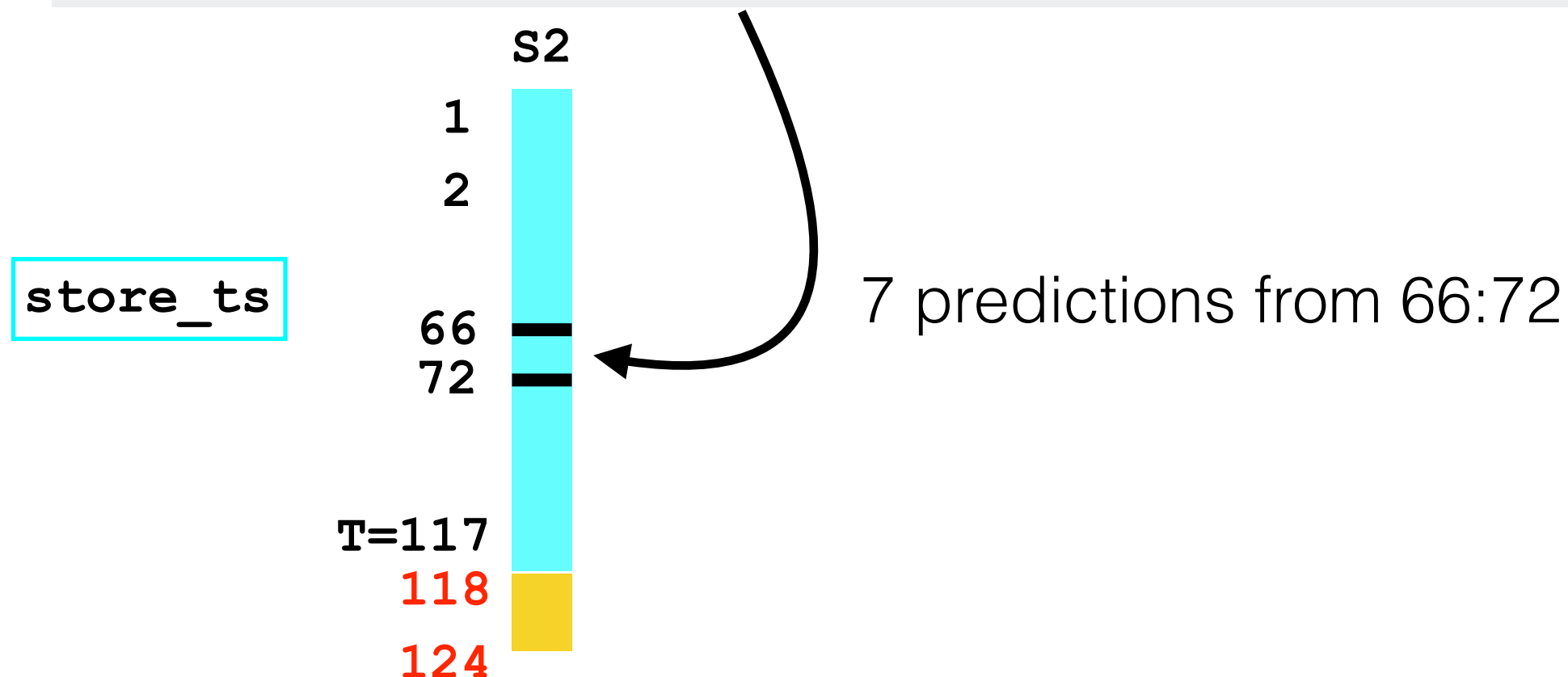


Focus on store 2; predict the next 7 wks

```
j=3 #Store 2
store_ts = ts(train_dept_ts[, j], frequency=52)
T = length(store_ts)
num_forecasts = 7
```

2. Seasonal Naive

```
snaive.pred = snaive(store_ts, num_forecasts)$mean
last.year.index = (T-52)+(1:num_forecasts)
store_ts[last.year.index]
snaive.pred
```



Focus on store 2; predict the next 7 wks

```
j=3 #Store 2
store_ts = ts(train_dept_ts[, j], frequency=52)
T = length(store_ts)
num_forecasts = 7
```

3. TSLM

```
myfit = tslm(store_ts ~ season + trend)
tslm.pred = forecast(myfit, h = num_forecasts)$mean
tslm.pred
```

	S2	Season	Trend
1		1	0
2		2	0
66		14	1
T=117		13	2
118		14	2
124		20	2

Linear model

Sales ~ Season + Trend

Season: categorical with 52 levels

Trend: numerical

3. TSLM

Linear model

Sales ~ Season + Trend

Season: categorical with 52 levels

Trend: numerical

```
myfit = tslm(store_ts ~ season + trend)
tslm.pred = forecast(myfit, h = num_forecasts)$mean
```

```
tmpdata = data.frame(
  Y = as.numeric(store_ts),
  season = as.factor((1:T) %% 52),
  trend = trunc((1:T)/52)
)
myfit = lm(Y ~ season + trend, tmpdata)

testindex = (T+1):(T+num_forecasts)
testdata = data.frame(
  season = as.factor(testindex %% 52),
  trend = trunc(testindex/52)
)

predict(myfit, newdata = testdata)
tslm.predtslm.pred
```