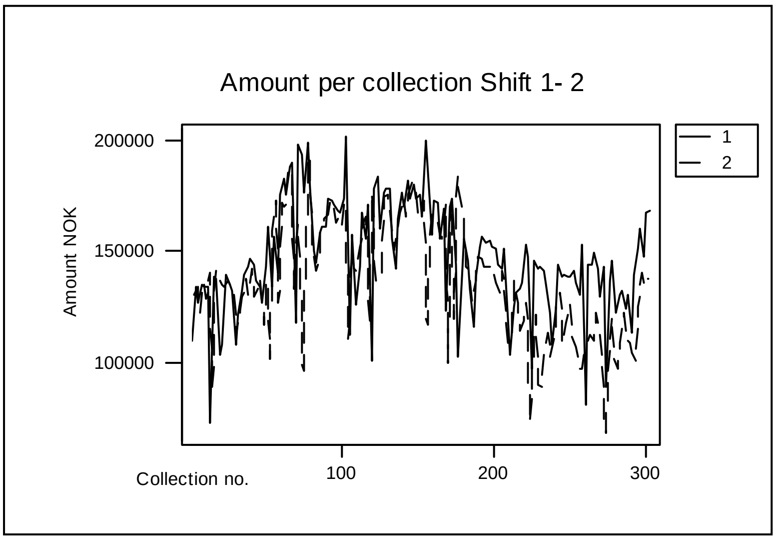
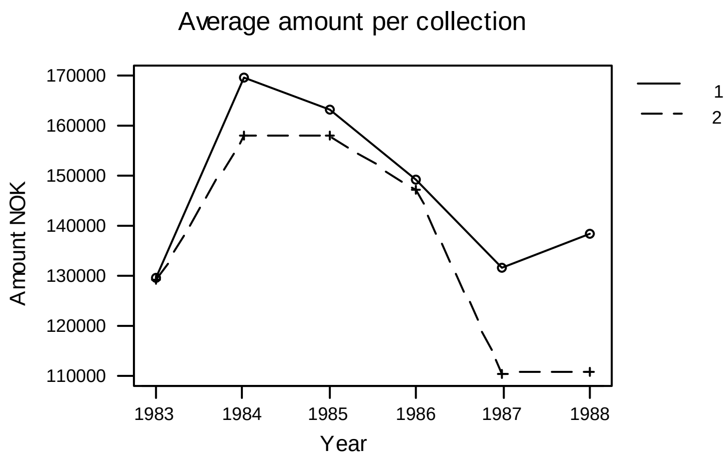
City Parking – Solution

A plot of average amount collected weekly for each shift follows.



We see that Shift 2 is below Shift 1 most of the time, with large differences for the latter collections. This becomes more apparent if we could zoom in parts of the plot. Alternatively we may plot the averages for each consecutive year for each shift:



We see that the yearly averages of Shift 2 are considerably below that of Shift 1 in 1984 and the last two years 1987 and 1988.

The differences are not likely to be due to chance alone, which will be confirmed by formal testing as follows:

Two-Sample T-Test and CI: Amount vs Shift (1)

```
Shift         N      Mean     StDev   SE Mean
1           149    147207     24344      1994
2           149    136362     26269      2152

Difference = mu (1) - mu (2)
Estimate for difference:  10846
95% CI for difference: (5071, 16620)
T-Test of difference = 0 (vs not =): T-Value = 3.70  P-Value = 0.000  DF = 294
```

We see that the hypothesis of equal mean amount for the two shifts is clearly rejected (P=0.000). One could alternatively perform a two-sample non-parametric test (Mann Whitney). This gives a similar negligible P-value. Looking closer at the data it is clear that observations at Christmas and Easter are outliers, but their removal does not affect P=0.000. Since they are few they do not matter much anyway.

The estimated mean difference of 10846 multiplied by 149 provides the estimate of the total amount embezzled of 1 616 048 mill. NOK. If we take the lower confidence limit 5071 literally, we can set a lower limit on the amount embezzled of about 750 000 with a 97.5% guarantee of catching the true amount above it. Is this justified or can we do better?

Concerning the assumptions for computing exact P-values and trustworthy confidence limits: Data for each shift over the range 1983-1988 hardly pass a common normality test (P-values for the Anderson-Darling statistic being P=0.070 and P=0.031 respectively). This is caused by the non-constant levels over time seen from the plot (also mentioned in the case description). This inflates the variances within groups, as well as the pooled variance, having the consequence of too small t-value and too wide confidence intervals. Note that the Mann-Whitney test is not really better justified. Although we get misleadingly wide confidence interval for its 97.5% guarantee, but the statistical significance is not ruined.

Let us therefore look at the data year for year (see plot above).

	Average per collection (NOK)				Fraud
Year	**Shift 1**	**Shift 2**	**Difference**	**N**	**estimates**
1983	**129 701**	**129 237**	**464**	**26**	**12 064**
1984	**169 416**	**157 839**	**11 577**	**26**	**301 002**
1985	**163 178**	**158 018**	**5 160**	**26**	**134 160**
1986	**149 285**	**147 235**	**2 050**	**24**	**49 200**
1987	**131 690**	**110 726**	**20 964**	**26**	**545 064**
1988	**138 449**	**111 091**	**27 358**	**21**	**574 518**
				Totalling NOK	**1 616 008**

Note. Two weeks of 1986 for Shift 2 omitted (no collection)

We may handle the holiday weeks separately, but this will not affect the estimates very much. We see that the total estimate obtained by aggregation over the years is about the same as above.

We now look at the t-tests separately for each year, after having removed observations for Christmas and Easter, as well as a period of strike in 1986, see computer output at the end.

The following conclusions are obtained: The hypothesis of equal mean amount for the two shifts is clearly rejected for the last two years (P=0.000), but not for any of the others at 5% significance level. For the second year (1984) P=0.090 (two-sided), so it is rejected on 10% level, but not on the 5% level. Whether the context justifies using the one-sided P=0.045 should be discussed (we think not). For the separate years the normality tests are passed, except for 1984 Shift 2 and 1987 Shift 1.

The corresponding, hopefully more realistic, confidence limits on the total amount may also be obtained by aggregation over years. Assuming independence between years we can obtain standard error of the total by taking the square root of the sum of squares of the standard error for each year, weighed by the number of Shift 2 weeks in that year. The standard error for each year, typically computed by pooling sum of squares deviation for each shift may be recovered from standard computer output. This computation gives a standard error of about 56 000 which gives a more realistic lower limit of 1 504 000 with about 97.5% guarantee of catching the true amount above it.

Although an improvement over the first analysis, this analysis assumes constant levels within each year, but we may have a seasonal pattern. There are different ways to overcome this. One is as follows: Create a sequence of "matched pairs" from subsequent amounts, collected by different shifts. Then analyze the differences between the amounts within each pair. This allows the level of parking income to vary over time. However, formal inference statements now require that the expected amounts taken away are approximately constant over time. This may of course a questionable assumption, but less so than the ones taken for the analysis above. Note that the pairing can be done two ways; pair a Shift 2 observation with its forward or backward neighbour. Analysis will show that the average of these differences does not differ much. and are about 11 000 with corresponding standard error of about 2 000. Projecting this to the 149 collections of Shift 2 gives the total estimate of about NOK 1 630 000, not much different from the take away estimate above. The corresponding standard error is about 25 000. Taken together this gives a lower limit of 1 580 000 with a 97.5% guarantee. In any case a conservative claim is that the take away is at least NOK 1.5 mill. (5 standard error down the tail).

<u>Note</u>. Autocorrelation in the differences may shrink the computed standard error and thus give a unrealistic high lower limit. This can be checked and does not seem to be the case.

A summary of the results so far (rounded to nearest thousand)

Analysis	Estimated mean amount taken	Lower limit for 97.5% guarantee
Overall (very naïve)	1 616 000	756 000
Yearly (naïve)	1 616 000	1 504 000
Matching pairs	1 657 000	1 608 000

The calculations above are admittedly crude, and may be improved by even more sophisticated methods. However, they are probably sufficient for the intended purpose. In practice one could defend beyond any doubt an amount of at least 1.5 mill. NOK, which is 5 times (instead of 2) the standard error down from the matching pair estimate.

It is felt that a two-factor analysis of variance will not provide new insight. As expected the computer output (3) shows a highly significant difference between the shifts after the differences between years are accounted for, and it also shows a significant interaction between shift and year, i.e. the shift differences are not uniform over the years. Looking at the residuals they fail the normality test, mainly due to a long left tail for both shifts, which may be partly due to some holidays not accounted for. We may want to perform a non-parametric test using shift as treatment and year as block. Common software may not include this, but test results will not differ anyway. The parametric ANOVA-model may also be basis for estimation of the total difference and corresponding confidence limits. Note however that the model assumes constant within year means, and is not likely to provide the kind of narrow limits as the “matched pair” approach.

Computer output

Two-Sample T-Test and CI: Amount vs Shift (1)

```
Shift         N      Mean     StDev   SE Mean
1           149    147207     24344      1994
2           149    136362     26269      2152

Difference = mu (1) - mu (2)
Estimate for difference:  10846
95% CI for difference: (5071, 16620)
T-Test of difference = 0 (vs not =): T-Value = 3.70  P-Value = 0.000  DF = 294
```

Two-Sample T-Test and CI: Amount83 versus Shift83 (2a)

```
Shift83      N      Mean     StDev   SE Mean
1           24    130789     12283      2507
2           24    132120      7420      1515

Difference = mu (1) - mu (2)
Estimate for difference:  -1330
95% CI for difference: (-7265, 4605)
T-Test of difference = 0 (vs not =): T-Value = -0.45  P-Value = 0.652  DF = 37
```

Two-Sample T-Test and CI: Amount84 versus Shift84 (2b)

```
Shift84      N      Mean     StDev   SE Mean
1           24    170256     17334      3538
2           24    160869     20117      4106

Difference = mu (1) - mu (2)
Estimate for difference:  9388
95% CI for difference: (-1530, 20305)
T-Test of difference = 0 (vs not =): T-Value = 1.73  P-Value = 0.090  DF = 45
```

Two-Sample T-Test and CI: Amount85 versus Shift85 (2c)

```
Shift85      N      Mean     StDev   SE Mean
1           24    164263     18052      3685
2           24    161456     14162      2891

Difference = mu (1) - mu (2)
Estimate for difference:  2807
95% CI for difference: (-6638, 12252)
T-Test of difference = 0 (vs not =): T-Value = 0.60  P-Value = 0.552  DF = 43
```

Two-Sample T-Test and CI: Amount86 versus Shift86 (2d)

```
Shift86      N      Mean     StDev   SE Mean
1           22    150393     17564      3745
2           23    149296     15777      3290

Difference = mu (1) - mu (2)
Estimate for difference:  1097
95% CI for difference: (-8962, 11156)
T-Test of difference = 0 (vs not =): T-Value = 0.22  P-Value = 0.827  DF = 42
```

Two-Sample T-Test and CI: Amount87 versus Shift87 (2e)

```
Shift87      N      Mean     StDev   SE Mean
1           23    134507     11814      2463
2           24    112632     12806      2614

Difference = mu (1) - mu (2)
Estimate for difference:  21876
95% CI for difference: (14637, 29114)
T-Test of difference = 0 (vs not =): T-Value = 6.09  P-Value = 0.000  DF = 44
```

Two-Sample T-Test and CI: Amount88 versus Shift88 (2f)

```
Shift88      N      Mean     StDev   SE Mean
1           21    138449     18069      3943
2           21    111091     16428      3585

Difference = mu (1) - mu (2)
Estimate for difference:  27358
95% CI for difference: (16580, 38137)
T-Test of difference = 0 (vs not =): T-Value = 5.13  P-Value = 0.000  DF = 39
```

General Linear Model: Amount_ versus Year_; Shift_ (3)

```
Factor     Type Levels Values
Year_     fixed      6 1983 1984 1985 1986 1987 1988
Shift_    fixed      2 1 2

Analysis of Variance for Amount_, using Adjusted SS for Tests

Source         DF     Seq SS     Adj SS     Adj MS       F      P
Year_           5 7.9792E+10 7.9835E+10 1.5967E+10   73.61  0.000
Shift_          1 7437007598 7955625725 7955625725   36.68  0.000
Year_*Shift_    5 7697074591 7697074591 1539414918    7.10  0.000
Error         264 5.7265E+10 5.7265E+10  216913443
Total         275 1.5219E+11

Term             Coef   SE Coef        T      P
Constant       143647       889   161.62  0.000
Shift_
1              5382.7     888.8     6.06  0.000
Year_
1983           -12192      1950    -6.25  0.000
1984            21916      1950    11.24  0.000
1985            19213      1950     9.85  0.000
1986             7332      2021     3.63  0.000
1987           -19693      1950   -10.10  0.000
Shift_*Year_
1      1983     -6048      1950    -3.10  0.002
1      1984      -689      1950    -0.35  0.724
1      1985     -3979      1950    -2.04  0.042
1      1986     -3700      2021    -1.83  0.068
1      1987      5939      1950     3.05  0.003
```

Two-Sample T-Test and CI: Deviation from trend versus Shift (4)

```
Shift         N      Mean     StDev   SE Mean
1           147      4374     15190      1253
2           147     -4317     14531      1199

Difference = mu (1) - mu (2)
Estimate for difference:  8691
95% CI for difference: (5279, 12104)
T-Test of difference = 0 (vs not =): T-Value = 5.01  P-Value = 0.000  DF = 291
```

Descriptive Statistics: Difference within pairs by Shift (5)

```
Variable   Shift              N          N*       Mean       StDev      SE Mean
DIFF       1                148           4      10968       22397         1841
           2                146           5     -11119       24686         2043
```