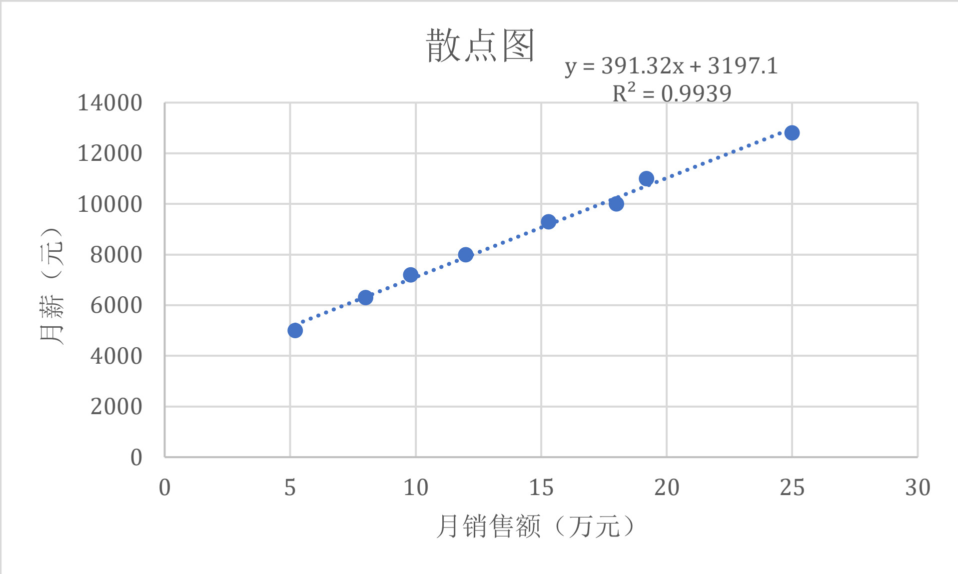
第 1 题：

（1）（2）



$$\bar{x} = 14.0625，\bar{y} = 8700.0$$

$$\hat{y} = \beta_1 x + \beta_0$$

$$\beta_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = 391.3175$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} = 8700 - 391.3175 \times 14.0625 = 3197.0971$$

$$R^2 = 1 - \frac{ss_{res}}{ss_{t_0t}}$$

$$ss_{t_0t} = \Sigma(y_i - \bar{y})^2 = 46340000.0$$

$$ss_{res} = \Sigma(y_i - \hat{y})^2 = 281926.7511$$

$$R^2 = 1 - \frac{281926.7511}{46340000.0} = 0.9939$$

与 excel 计算结果相吻合

（3）

平均绝对误差（MAE）：170.1085

均方误差（MSE）：35240.7189

这两个指标都比较小，进一步说明线性回归模型拟合效果良好。

第 3 题：

计算每个类别的打分

$$z_j = \omega_j^T x_1 + b_j$$

Softmax 输出

$$\hat{y}_i = \frac{e^{z_j}}{\sum_{k=1}^{l} e^{z_k}}$$

设$y_1 = c$，即样本属于第 c 类

$$L = -log(\hat{y}_c)$$

对于每个类别 j：

权重ω_j的梯度为

$$\frac{\partial L}{\partial \omega_j} = (\hat{y}_j - \delta_{jc})x_1$$

偏置b_j的梯度为

$$\frac{\partial L}{\partial b_j} = \hat{y}_j - \delta_{jc}$$

其中$\delta_{jc} = \begin{cases} 1，if\ j = c \\ 0，other \end{cases}$

使用学习率 α，梯度下降更新规则如下：

对于每个类别 j：

$$\omega_j \leftarrow \omega_j - \alpha \cdot \frac{\partial L}{\partial \omega_j} = \omega_j - \alpha(\hat{y}_j - \delta_{jc})x_1$$

$$b_j \leftarrow b_j - \alpha \cdot \frac{\partial L}{\partial b_j} = b_j - \alpha(\hat{y}_j - \delta_{jc})$$