Task 4

Basically, RBF kernel-value can be interpreted as similarity value. Similarity will be $\in [0,1]$ where 1 denotes maximum similarity. Having m datapoints, there are m similarity values for this point, one for each other datapoint plus the similarity to itsself. Since we can have an infinite number of points, the resulting feature space can also be infinite.

Proof:

For simplicity, we choose σ such that $2\sigma^2 = 1$

then: $k(x,x') = e^{-\|x-x'\|^2}$

Let x and x' be $\in \mathbb{R}^2$ with $x = (x_1, x_2)$ and $x' = (x_1', x_2')$

$$k(x,x') = e^{-((x_1-x_1')^2 + (x_2-x_2')^2)}$$

$$= \exp\left(-\left(x_1^2 - 2x_1x_1' + x_1'^2 + x_2^2 - 2x_2x_2' + x_2'^2\right)\right)$$

$$= \exp\left(-x_1^2 + 2x_1x_1' - x_1'^2 - x_2^2 + 2x_2x_2' - x_2'^2\right)$$

$$= \exp\left(-\|x\|^2\right) \cdot \exp\left(-\|x'\|^2\right) \cdot \exp\left(2x_1x_1' + 2x_2x_2'\right)$$

$$= \exp\left(-\|x\|^2\right) \cdot \exp\left(-\|x'\|^2\right) \cdot \exp\left(2x^Tx'\right)$$

Now $2x^Tx'$ already looks like a polynomial kernel of degree 1 we use taylor expansion on $e^{2x^Tx'}$

$$= \exp\left(-\|x\|^2\right)\exp\left(-\|x'\|^2\right) \cdot \sum_{n=0}^{\infty} \frac{(2x^Tx)^n}{n!}$$

$\sum_{n=0}^{\infty} \frac{(2x^Tx)^n}{n!}$ means that we are adding

infinitely many polynomial kernels with increasing (up to infinite) degree.
When adding two polynomial kernels, the resulting feature space has the sum of the ~~feature spaces~~ added kernels' feature spaces as dimension, thus the RBF kernel has an infinite feature space.