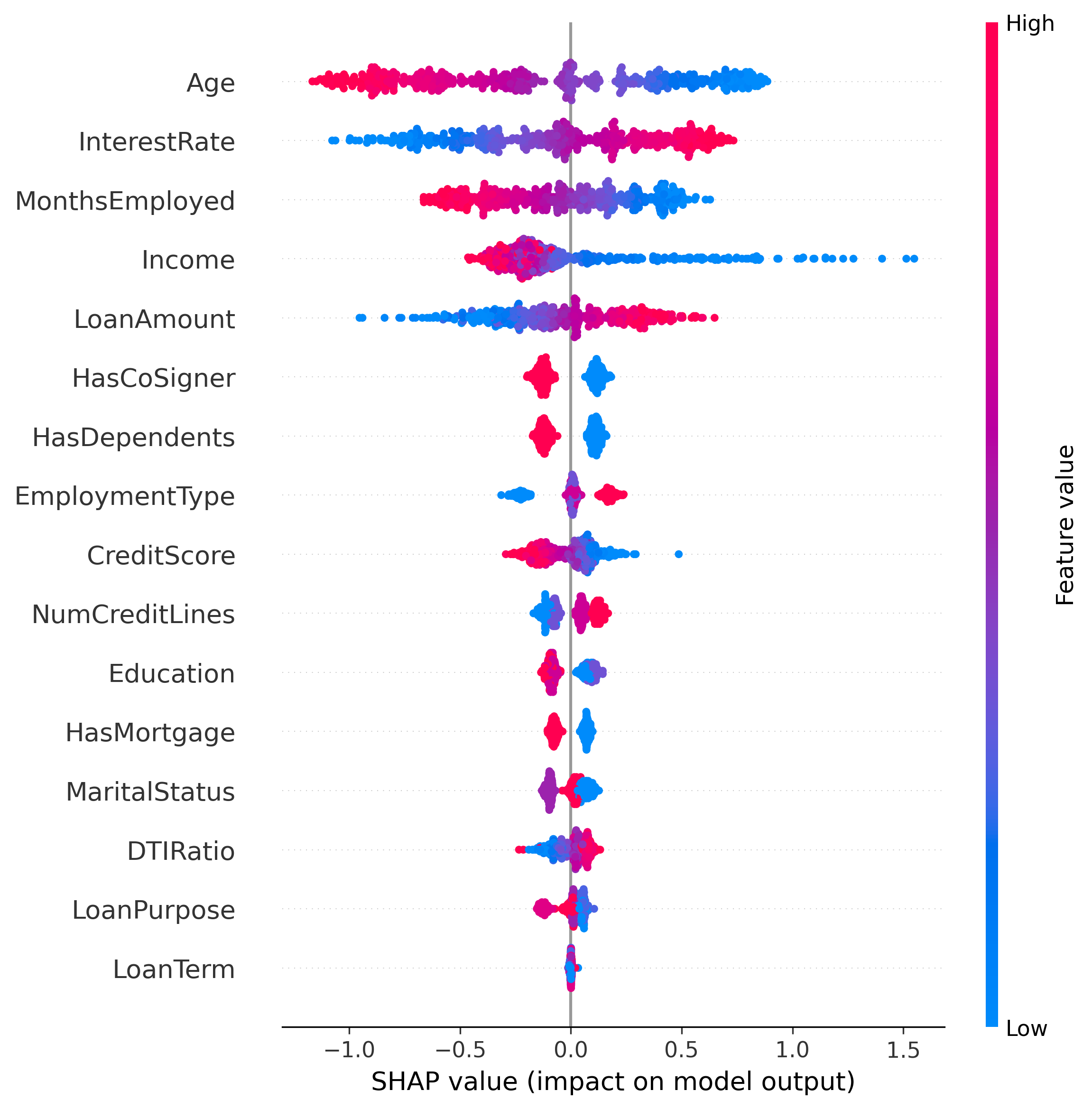
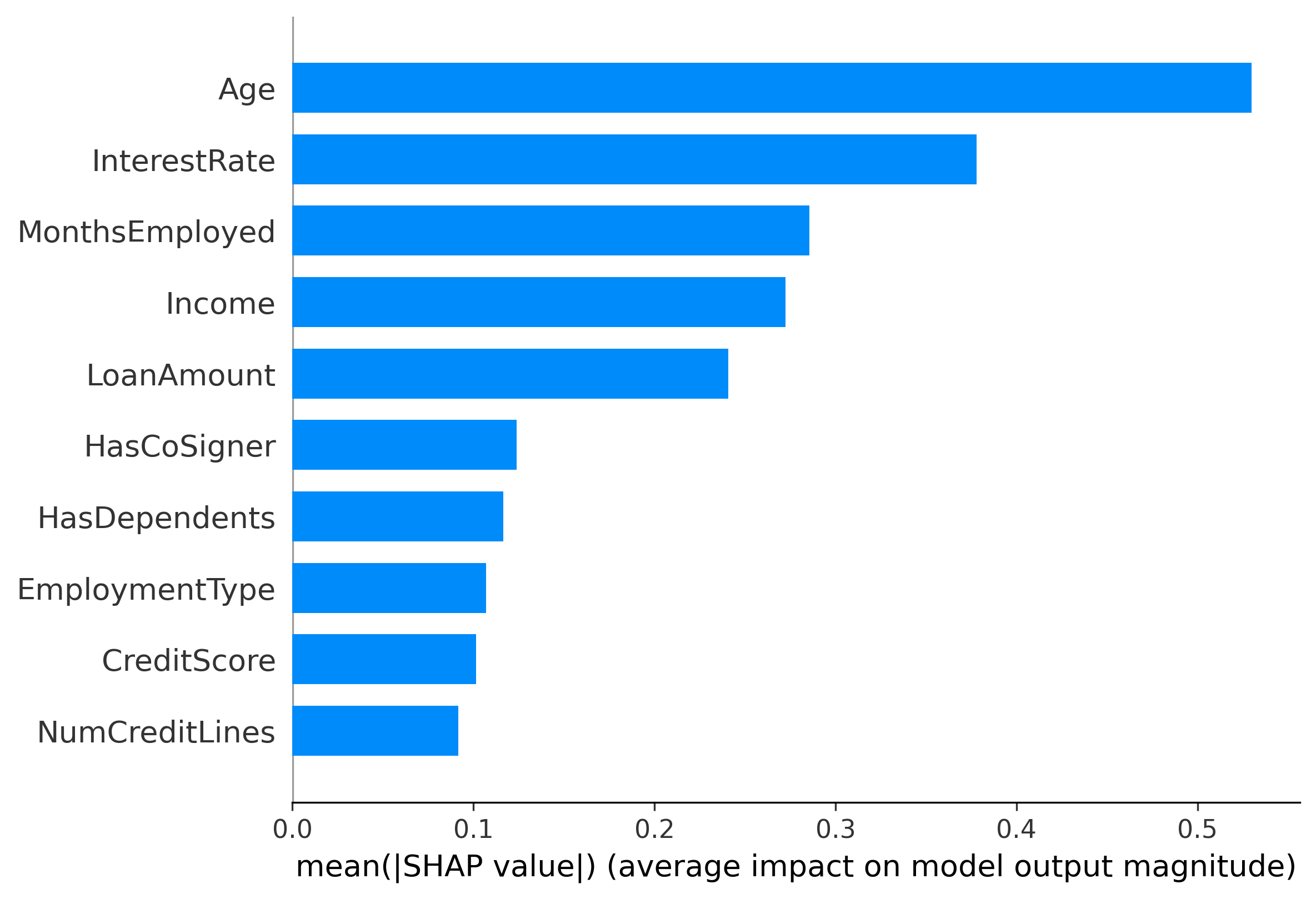
# SHAP Value Analysis - Loan Default Prediction Model

## Overview

We used SHAP (SHapley Additive exPlanations) values to interpret our XGBoost loan default prediction model. SHAP values help us understand how each feature impacts the model's output for both individual samples and overall trends.

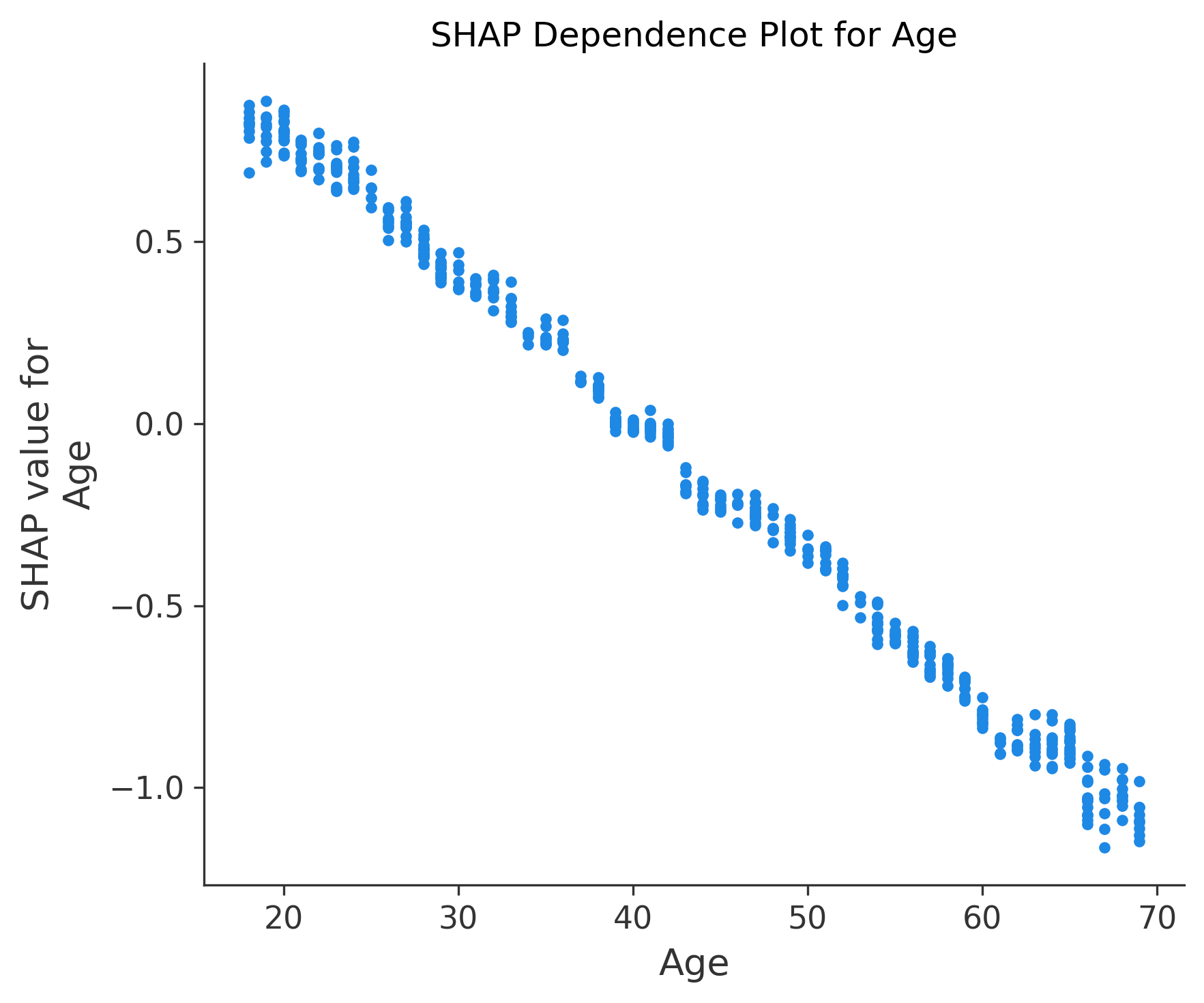


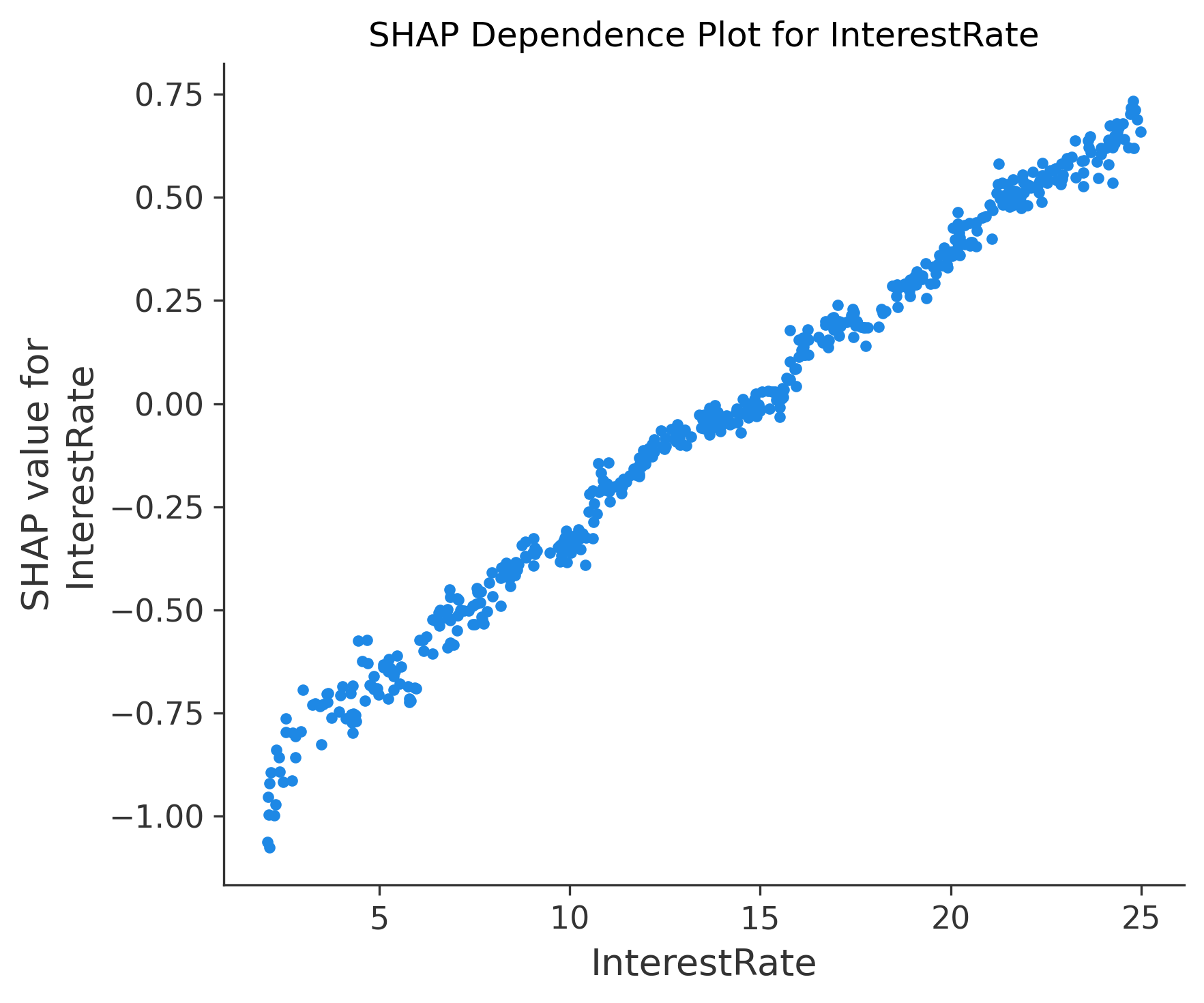


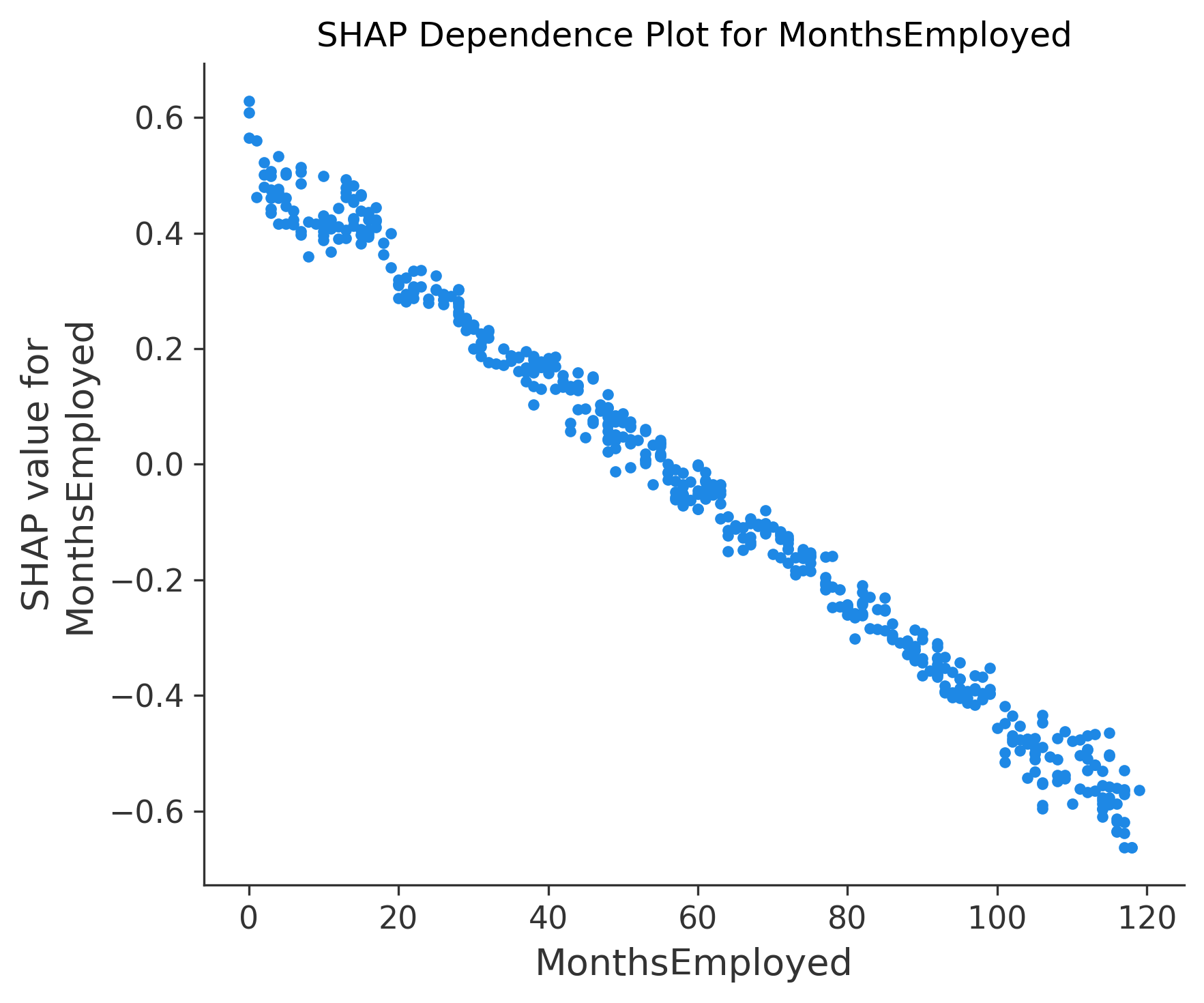
## Top Important Features

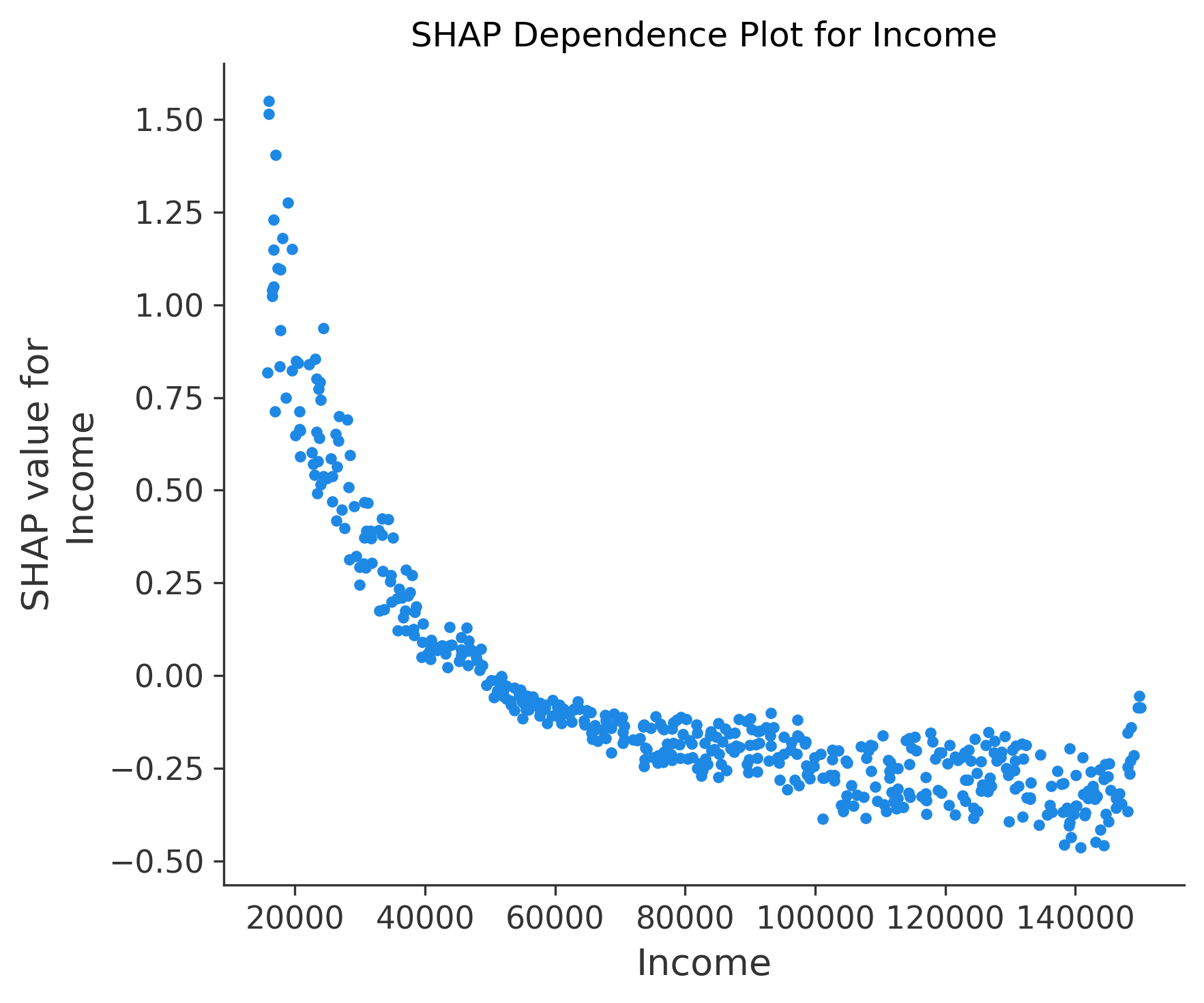
Based on the SHAP summary and feature importance plots, the following features have the greatest influence on the model's predictions:  
1. Age – Younger borrowers have higher SHAP values, indicating increased default risk.  
2. Interest Rate – Higher interest rates strongly increase the probability of default.  
3. Months Employed – Longer employment history is associated with lower default risk.  
4. Income – Lower income significantly contributes to a higher likelihood of default.  
5. Loan Amount – Larger loan amounts are correlated with higher risk.  
6. HasCoSigner, HasDependents, EmploymentType – also have modest impact.

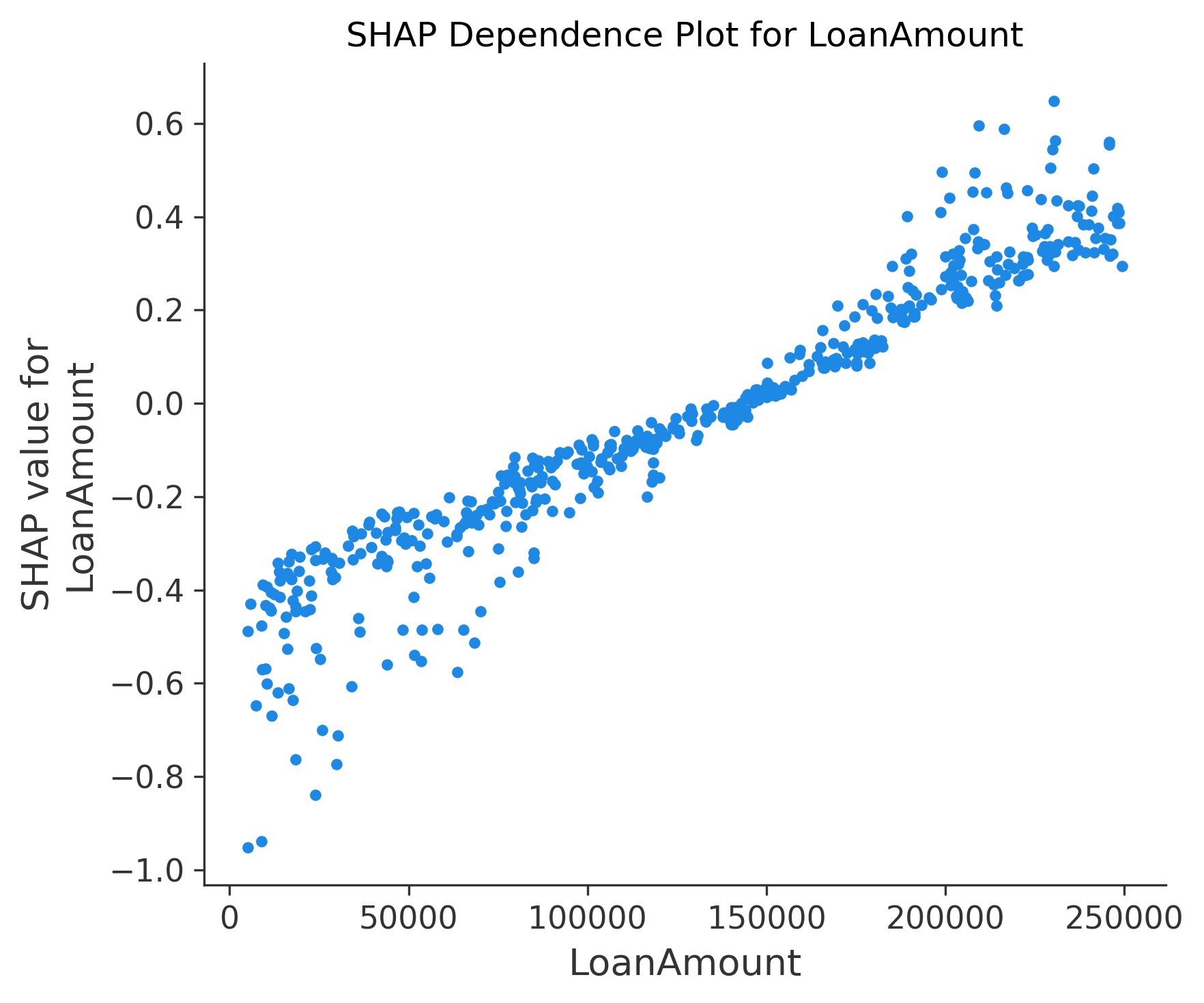
## SHAP Dependence Plots Highlights

  
- Age: A clear negative relationship — as age increases, SHAP value decreases → older individuals are less likely to default.

  
- Interest Rate: Strong positive effect — higher rates push SHAP values higher, increasing predicted risk.

  
- Months Employed: The longer someone has been employed, the less risky they are (SHAP values drop).

  
- Income: Lower income corresponds to higher SHAP values and greater predicted risk.

  
- Loan Amount: Larger loan amounts tend to push risk higher, especially above ~$150,000.  
  
These patterns align with real-world intuition: younger, less stable, lower-income borrowers with larger loans and higher rates are more likely to default.